

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 19:05:42 ; Search time 51 Seconds
(without alignments)
626.037 Million cell updates/sec

Title: US-09-756-541-1
Perfect score: 561
Sequence: 1 FYKVEALVPRRIQQVSSAL.....VLPVSDVIRVTRGEGKAE 113

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 6
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 29Jan04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	561	100.0	113	4	AAB69495 Arabidops
2	509	90.7	113	4	AAB69496 Ricinus c
3	312.5	55.7	112	4	AAB69503 Synchoco
4	308.5	55.0	121	6	Abp79052 N. gonorr
5	287.5	51.2	112	4	AAB69501 Azospiril
6	284.5	50.7	112	4	AAB69497 Klebsiell
7	284.5	50.7	112	4	AAB69498 Escherich
8	278.5	49.6	112	4	AAG93226 C. glutam
9	264.5	47.1	112	4	AAB69502 Rhodobact
10	260.5	46.4	124	6	Ada34150 Acinetoba
11	257	45.8	111	4	AAB69499 Rhizobium
12	256.5	45.7	160	4	Abg24254 Novel hum
13	256.5	45.6	111	4	AAB69500 Bradyrhiz
14	246.5	43.9	89	4	Abp79181 Corynebact
15	225	40.1	112	5	ABBS4937 Lactococc
16	217.5	38.8	91	5	ABP03233 Human ORF
17	205.5	36.6	121	5	ABE49279 Listeria
18	203	36.2	243	4	Abg24253 Novel hum
19	203	36.2	579	4	Abg24250 Novel hum
20	183	32.6	111	4	AAB69505 Methanoco
21	162.5	29.0	1285	4	Abg26219 Novel hum
22	155.5	27.7	105	4	AAB69504 Methanoco
23	125	22.3	736	4	Abg24252 Novel hum
24	80.5	14.3	1036	6	Abu49254 Protein e
25	74.5	13.3	154	2	AAR05799 PBOMP-2 g

ALIGNMENTS

RESULT 1
AAB69495
ID AAB69495 standard; protein; 113 AA.

XX AAB69495;
DT 23-APR-2001 (first entry)
DE Arabidopsis thaliana PII protein.
KW Arabidopsis thaliana; PII; plant nitrogen regulatory gene; P-II;
KW nitrogen assimilation; transgenic plant; herbicide screening.
XX Arabidopsis thaliana.

XX US6177275-B1.
XX 23-JAN-2001.
XX 23-JUL-1997; 97US-00899330.
XX 24-JUL-1996; 96US-0022328P.
XX (UUNY) UNIV NEW YORK STATE.
XX Coruzzi GM, Lam H, Hsieh M;
XX WPI; 2001-158572/16.
XX N-PSDB; AAF58581, AAF58583.

XX Novel P-II genes capable of regulating plant nitrogen assimilation,
XX useful for transgenic plant production, and as probes for isolating
XX additional genomic clones having P-II gene promoters.
XX Claim 1; Fig 1; 35pp; English.
XX The present sequence is encoded by a nitrogen regulatory PII gene. Novel
XX plant PII (also called P-II) nucleotide sequences have been isolated.
XX They are useful for regulating nitrogen assimilation in plants, and in
XX transgenic plant production. They are also used to engineer organisms
XX that overexpress wild-type or mutant P-II regulatory proteins. P-II
XX proteins are useful for in vitro screening of herbicides. P-II
XX nucleotides may be used as probes for isolating additional genomic clones
XX with the promoters of P-II genes. P-II promoters are light- and/or
XX sucrose-inducible, and are suitable for genetic engineering of plants
XX Sequence 113 AA;

Abg21490 Novel hum
Abu25143 Protein e
Aay31990 Chimeric
Adb12369 Allostococ
Aar14528 C-termina
Aar14019 Insectici
Aar75409 B. sphaer
Aau02093 Bacillus
Aau02094 Bacillus
Aau02092 Bacillus
Aar50955 Bacillus
Abp66214 Bifidobac
Aaw73375 S. aureus
Abm71198 Staphyloc
Abu06030 N. mening
Abu35633 Protein e
Abm65566 Propionib
Aau44161 Propionib
Abm40680 Propionib
Abp76681 Streptomy

Query Match 100.0%; Score 561; DB 4; Length 113;
 Best Local Similarity 100.0%; Pred. No. 1.8e-62;
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FYKVEAIVPWRIQOVSSALLKIGIRGVTVDVGRFGAOGGSTERHGGSEFSEDKFVAKV 60
 DB 1 FYKVEAIVPWRIQOVSSALLKIGIRGVTVDVGRFGAOGGSTERHGGSEFSEDKFVAKV 60

QY 61 KMEIVVKQOVESVINTIIEGARTGEIGDGKIFVLPSVDVIRVTRTGERGEKAE 113
 DB 61 KMEIVVKQOVESVINTIIEGARTGEIGDGKIFVLPSVDVIRVTRTGERGEKAE 113

RESULT 2
 AAB69496
 ID AAB69496 standard; protein; 113 AA.
 AC AAB69496;
 XX
 XX 23-APR-2001 (first entry)
 DT
 XX Ricinus communis PII protein.
 DE
 XX Castor bean; PII; plant nitrogen regulatory gene; P-II;
 KW nitrogen assimilation; transgenic plant; herbicide screening.
 KW
 XX Ricinus communis.
 OS
 XX US6177275-B1.
 PN
 XX 23-JAN-2001.
 PD
 XX 23-JUL-1997; 97US-00899330.
 PF
 XX 24-JUL-1996; 96US-0022328P.
 PR
 XX (UUNY) UNIV NEW YORK STATE.
 PA
 XX Coruzzi GM, Lam H, Hsieh M;
 PI WPI; 2001-158572/16.
 DR
 XX Novel P-II genes capable of regulating plant nitrogen assimilation,
 PT useful for transgenic plant production, and as probes for isolating
 PT additional genomic clones having P-II gene promoters.
 XX
 PS Claim 1; Fig 1; 35pp; English.
 SQ

The present sequence is encoded by a nitrogen regulatory PII gene. Novel plant PII (also called P-II) nucleotide sequences have been isolated. They are useful for regulating nitrogen assimilation in plants, and in transgenic plant production. They are also used to engineer organisms that overexpress wild-type or mutant P-II regulatory proteins. P-II proteins are useful for in vitro screening of herbicides. P-II nucleotides may be used as probes for isolating additional genomic clones with the promoters of P-II genes. P-II promoters are light- and/or sucrose-inducible, and are suitable for genetic engineering of plants

Query Match 90.7%; Score 509; DB 4; Length 113;
 Best Local Similarity 90.3%; Pred. No. 6.3e-56;
 Matches 102; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 FYKVEAIVPWRIQOVSSALLKIGIRGVTVDVGRFGAOGGSTERHGGSEFSEDKFVAKV 60
 DB 1 FYKVEAIVPWRVQVSSALLKIGIRGVTVDVGRFGAOGGSTERHGGSEFSEDKFVAKV 60

QY 61 KMEIVVKQOVESVINTIIEGARTGEIGDGKIFVLPSVDVIRVTRTGERGEKAE 113
 DB 61 KMEIVVKQOVEDVTEKIEEARTGEIGDGKIFLLPSVDVIRVTRTGERGDKAE 113

RESULT 3
 AAB69503
 ID AAB69503 standard; protein; 112 AA.
 AC AAB69503;
 XX
 XX 23-APR-2001 (first entry)
 DT
 XX Synechococcus strain PCC 7942 PII protein.
 DE
 XX Synechococcus; PII; plant nitrogen regulatory gene; P-II;
 KW nitrogen assimilation; transgenic plant; herbicide screening.
 KW
 XX Synechococcus sp.
 OS
 XX US6177275-B1.
 PN
 XX 23-JAN-2001.
 PD
 XX 23-JUL-1997; 97US-00899330.
 PF
 XX 24-JUL-1996; 96US-0022328P.
 PR
 XX (UUNY) UNIV NEW YORK STATE.
 PA
 XX Coruzzi GM, Lam H, Hsieh M;
 PI WPI; 2001-158572/16.
 DR
 XX Novel P-II genes capable of regulating plant nitrogen assimilation,
 PT useful for transgenic plant production, and as probes for isolating
 PT additional genomic clones having P-II gene promoters.
 XX
 PS Example; Fig 1; 35pp; English.
 SQ

The present sequence is encoded by a nitrogen regulatory PII gene. Novel plant PII (also called P-II) nucleotide sequences have been isolated. They are useful for regulating nitrogen assimilation in plants, and in transgenic plant production. They are also used to engineer organisms that overexpress wild-type or mutant P-II regulatory proteins. P-II proteins are useful for in vitro screening of herbicides. P-II nucleotides may be used as probes for isolating additional genomic clones with the promoters of P-II genes. P-II promoters are light- and/or sucrose-inducible, and are suitable for genetic engineering of plants

Query Match 55.7%; Score 312.5; DB 4; Length 112;
 Best Local Similarity 56.4%; Pred. No. 3.7e-31;
 Matches 62; Conservative 23; Mismatches 24; Indels 1; Gaps 1;

QY 3 KVEAIVPWRIQOVSSALLKIGIRGVTVDVGRFGAOGGSTERHGGSEFSEDKFVAKV 62
 DB 3 KIEAIIIEPFKLDKIEKIALVNAVIGVMTVSEVRGFGKQGTQRYRGSEITVE-FLQKLL 61

QY 63 EIVVKQOVESVINTIIEGARTGEIGDGKIFVLPSVDVIRVTRTGERGEKA 112
 DB 62 EIVVEDAQVDIVDKIVAAARTGEIGDGKIFVSPVDQTIRITRTEKNADA 111

RESULT 4
 ABP79052
 ID ABP79052 standard; protein; 121 AA.
 AC ABP79052;
 XX
 XX 07-MAR-2003 (first entry)
 DT
 XX N. gonorrhoeae amino acid sequence SEQ ID 4634.
 DE
 XX Antibacterial; infection; vaccine; gene therapy.
 XX

OS Neisseria gonorrhoeae.
 XX
 XX WO200279243-A2.
 XX
 PD 10-OCT-2002.
 XX
 XX 12-FEB-2002; 2002WO-IB002069.
 XX
 XX 12-FEB-2001; 2001GB-00003424.
 XX
 XX (CHIR-) CHIRON SPA.
 XX
 XX Fontana MR, Pizza M, Masignani V, Monaci E;
 XX
 XX WPI; 2003-058415/05.
 DR N-PSDB; AB240022.
 XX
 XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a
 PT medicament for treating or preventing N. gonorrhoeae infection.
 XX
 XX Disclosure; Page 528; 815pp; English.
 XX
 CC The present invention relates to proteins from Neisseria gonorrhoeae.
 CC Also disclosed are the nucleic acid molecules encoding the proteins and
 CC antibodies that specifically bind to the proteins. The composition
 CC comprising the protein, nucleic acid or antibody is useful for the
 CC manufacture of a medicament for treating or preventing N. gonorrhoeae
 CC infection, this may be in the form of a vaccine or gene therapy.
 CC Sequences given in records ABP76736-ABP81046 represent nucleic acid
 CC molecules of the invention
 XX
 SQ Sequence 121 AA;
 Query Match 55.0%; Score 308.5; DB 6; Length 121;
 Best Local Similarity 55.3%; Pred. No. 1.3e-30;
 Matches 61; Conservative 24; Mismatches 24; Indels 1; Gaps 1;
 QY 3 KVEAIVRPWRIQQVSSALLKIGIRGTVSDVRGFGAGGSTERHGGSEDFKFAVKVM 62
 Db 12 KIEAIVKPKDDVRALCEIGITGVSEVKGFGKQKGTETLYRGAYVD-FLPKVKI 70
 QY 63 EIVVKKQDVESVINTIEGARTGEIGDKIFVLPSDVIRVIRVTGERGEKA 112
 Db 71 ELVLADDAVERAIDVIVEVARSKIGDKIFVLPSVIRVIRVTGERGEKA 120
 RESULT 5
 ID AAB69501 standard; protein; 112 AA.
 XX
 AC AAB69501;
 XX
 DT 11-SEP-2003 (revised)
 DT 23-APR-2001 (first entry)
 XX
 DE Azospirillum brasilense PII protein.
 XX
 XX Azospirillum brasilense; PII; plant nitrogen regulatory gene; P-II;
 KW nitrogen assimilation; transgenic plant; herbicide screening.
 XX
 OS Azospirillum brasilense.
 XX
 XX US6177275-B.
 XX
 PD 23-JAN-2001.
 XX
 XX 23-JUL-1997; 97US-00899330.
 XX
 XX 24-JUL-1996; 96US-0022328P.
 XX
 XX (UYN Y) UNIV NEW YORK STATE.
 XX
 XX Coruzzi GM, Lam H, Hsieh M;

XX WPI; 2001-158572/16.
 XX
 XX Novel P-II genes capable of regulating plant nitrogen assimilation,
 PT useful for transgenic plant production, and as probes for isolating
 PT additional genomic clones having P-II gene promoters.
 XX
 XX Example; Fig 1; 35pp; English.
 XX
 CC The present sequence is encoded by a nitrogen regulatory PII gene. Novel
 CC plant PII (also called P-II) nucleotide sequences have been isolated.
 CC They are useful for regulating nitrogen assimilation in plants, and in
 CC transgenic plant production. They are also used to engineer organisms
 CC that overexpress wild-type or mutant P-II regulatory proteins. P-II
 CC proteins are useful for in vitro screening of herbicides. P-II
 CC nucleotides may be used as probes for isolating additional genomic clones
 CC with the promoters of P-II genes. P-II promoters are light- and/or
 CC sucrose-inducible, and are suitable for genetic engineering of plants.
 CC (Updated on 11-SEP-2003 to standardise OS field)
 XX
 SQ Sequence 112 AA;
 Query Match 51.2%; Score 287.5; DB 4; Length 112;
 Best Local Similarity 49.1%; Pred. No. 5.3e-28;
 Matches 54; Conservative 27; Mismatches 28; Indels 1; Gaps 1;
 QY 3 KVEAIVRPWRIQQVSSALLKIGIRGTVSDVRGFGAGGSTERHGGSEDFKFAVKVM 62
 Db 3 KIEAIVKPKDDVRALCEIGITGVSEVKGFGKQKGTETLYRGAYVD-FLPKVKI 61
 QY 63 EIVVKKQDVESVINTIEGARTGEIGDKIFVLPSDVIRVIRVTGERGEKA 112
 Db 62 EVWEDSLVERAIEAIOQAATGRIQDKIFVTPVEVIRVTGEGDA 111
 RESULT 6
 ID AAB69497 standard; protein; 112 AA.
 XX
 AC AAB69497;
 XX
 DT 23-APR-2001 (first entry)
 XX
 DE Klebsiella pneumoniae PII protein.
 XX
 XX Klebsiella pneumoniae; PII; plant nitrogen regulatory gene; P-II;
 KW nitrogen assimilation; transgenic plant; herbicide screening.
 XX
 OS Klebsiella pneumoniae.
 XX
 XX US6177275-B1.
 XX
 PD 23-JAN-2001.
 XX
 XX 23-JUL-1997; 97US-00899330.
 XX
 XX 24-JUL-1996; 96US-0022328P.
 XX
 XX (UYN Y) UNIV NEW YORK STATE.
 XX
 XX Coruzzi GM, Lam H, Hsieh M;
 XX
 XX WPI; 2001-158572/16.
 XX
 XX Novel P-II genes capable of regulating plant nitrogen assimilation,
 PT useful for transgenic plant production, and as probes for isolating
 PT additional genomic clones having P-II gene promoters.
 XX
 XX Example; Fig 1; 35pp; English.
 XX
 CC The present sequence is encoded by a nitrogen regulatory PII gene. Novel
 CC plant PII (also called P-II) nucleotide sequences have been isolated.
 CC They are useful for regulating nitrogen assimilation in plants, and in

CC transgenic plant production. They are also used to engineer organisms
 CC that overexpress wild-type or mutant P-II regulatory proteins. P-II
 CC proteins are useful for in vitro screening of herbicides. P-II
 CC nucleotides may be used as probes for isolating additional genomic clones
 CC with the promoters of P-II genes. P-II promoters are light- and/or
 CC sucrose-inducible, and are suitable for genetic engineering of plants
 XX
 SQ Sequence 112 AA;

Query Match 50.7%; Score 284.5; DB 4; Length 112;
 Best Local Similarity 50.9%; Pred. No. 1.3e-27;
 Matches 56; Conservative 27; Mismatches 26; Indels 1; Gaps 1;

QY 3 KVEAIVRPMRIQOVSSALLKIGIRGTVSDVRGFGAQQGSTERHGGSEFSEDKFVAKVKM 62
 DB 3 KIDAILKPKFLDDVREALAEVGTMTVTBKFGKQKGTHTLYRGAAYMD-FLPKVKI 61

QY 63 EIVVKDQVESVINTIIEGARTGEIGDKIFVLPSDVIRVTGERGEKA 112
 DB 62 EIVVTDIVTCDVTIIRTAQTGKIGDKIFVFDVARVIRIRTGEDDAA 111

RESULT 7
 ID AAB69498 standard; protein; 112 AA.
 AC AAB69498;

DT 23-APR-2001 (first entry)
 XX Escherichia coli PII protein.
 DE Escherichia coli; PII; plant nitrogen regulatory gene; P-II;
 KW nitrogen assimilation; transgenic plant; herbicide screening.
 XX

OS Escherichia coli.
 XX
 PK US6177275-B1.
 XX

PD 23-JAN-2001.
 XX
 PF 23-JUL-1997; 97US-00899330.
 XX

PR 24-JUL-1996; 96US-0022328P.
 XX
 PA (UJNY) UNIV NEW YORK STATE.
 XX

PI Coruzzi GM, Lam H, Hsieh M;
 XX
 DR WPI; 2001-158572/16.
 XX

PT Novel P-II genes capable of regulating plant nitrogen assimilation,
 PT useful for transgenic plant production, and as probes for isolating
 PT additional genomic clones having P-II gene promoters.
 XX

PS Example; Fig 1; 35pp; English.
 XX

CC The present sequence is encoded by a nitrogen regulatory PII gene. Novel
 CC plant PII (also called P-II) nucleotide sequences have been isolated.
 CC They are useful for regulating nitrogen assimilation in plants, and in
 CC transgenic plant production. They are also used to engineer organisms
 CC that overexpress wild-type or mutant P-II regulatory proteins. P-II
 CC proteins are useful for in vitro screening of herbicides. P-II
 CC nucleotides may be used as probes for isolating additional genomic clones
 CC with the promoters of P-II genes. P-II promoters are light- and/or
 CC sucrose-inducible, and are suitable for genetic engineering of plants
 XX
 SQ Sequence 112 AA;

Query Match 50.7%; Score 284.5; DB 4; Length 112;
 Best Local Similarity 50.9%; Pred. No. 1.3e-27;
 Matches 56; Conservative 27; Mismatches 26; Indels 1; Gaps 1;

QY 3 KVEAIVRPMRIQOVSSALLKIGIRGTVSDVRGFGAQQGSTERHGGSEFSEDKFVAKVKM 62
 DB 3 KIDAILKPKFLDDVREALAEVGTMTVTBKFGKQKGTHTLYRGAAYMD-FLPKVKI 61
 QY 63 EIVVKDQVESVINTIIEGARTGEIGDKIFVLPSDVIRVTGERGEKA 112
 DB 62 EIVVTDIVTCDVTIIRTAQTGKIGDKIFVFDVARVIRIRTGEDDAA 111

RESULT 8
 ID AAG93226 standard; protein; 112 AA.
 XX AAG93226;

AC AAG93226;
 DT 26-SEP-2001 (first entry)
 XX

DE C glutamicum protein fragment SEQ ID NO: 6980.

KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis.
 XX

OS Corynebacterium glutamicum.
 XX

PN EPI108790-A2.
 XX

PD 20-JUN-2001.
 XX

PF 18-DEC-2000; 2000EP-00127688.
 XX

PR 16-DEC-1999; 99JP-00377484.
 XX

PR 07-APR-2000; 2000JP-00159162.
 XX

PR 03-AUG-2000; 2000JP-00280988.
 XX

FA (KYOW) KYOWA HAKKO KOGYO KK.
 XX

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX

DR WPI; 2001-376931/40.
 XX

DR N-ESDE; AAH68445.
 XX

PT Novel polynucleotides derived from Coryneform bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX

PS Claim 29; SEQ ID NO 6980; 246pp + Sequence Listing; English.
 XX

CC The present invention provides a number of nucleotide and protein
 CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of coryneform bacterium, measuring expression amount and analysing
 CC the expression profile or expression pattern of a gene derived from
 CC Coryneform bacterium, and identifying a homologue of a gene derived from
 CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described in the
 CC exemplification of the invention. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from the European Patent Office
 XX

SQ Sequence 112 AA;

Query Match 49.6%; Score 278.5; DB 4; Length 112;
 Best Local Similarity 46.8%; Pred. No. 7.2e-27;
 Matches 51; Conservative 30; Mismatches 27; Indels 1; Gaps 1;

QY 4 VEAIVRPMRIQOVSSALLKIGIRGTVSDVRGFGAQQGSTERHGGSEFSEDKFVAKVKM 63
 DB 4 ITAIVKPTLTDIKDALEQAGVQGTMTVTETQFGQKQKGTHTLYRGAAYMD-FVLPKVKIE 62

QY 64 IWKVKDQVESVINTIIEGARTGEIGDKIFVLPSDVIRVTGERGEKA 112
 DB 62 EIVVTDIVTCDVTIIRTAQTGKIGDKIFVFDVARVIRIRTGEDDAA 111

Db 63 VIISDAQAEVINIIVTARTKVGDKGKVMNIBELVVRVTRGERGEAA 111

RESULT 9
 AAB69502
 ID AAB69502 standard; protein; 112 AA.
 XX
 AC AAB69502;
 XX
 DT 23-APR-2001 (first entry)
 XX
 DE Rhodobacter capsulatus PII protein.
 XX
 KW Rhodobacter capsulatus; PII; plant nitrogen regulatory gene; P-II;
 KW nitrogen assimilation; transgenic plant; herbicide screening.
 XX
 OS Rhodobacter capsulatus.
 XX
 PN US6177275-B1.
 XX
 PD 23-JAN-2001.
 XX
 PF 23-JUL-1997; 97US-00899330.
 XX
 PR 24-JUL-1996; 96US-0022328P.
 XX
 PA (UNY) UNIV NEW YORK STATE.
 XX
 PI Coruzzi GM, Lam H, Hsieh M;
 XX
 DR WPI; 2001-158572/16.
 XX
 PT Novel P-II genes capable of regulating plant nitrogen assimilation,
 PT useful for transgenic plant production, and as probes for isolating
 PT additional genomic clones having P-II gene promoters.
 XX
 PS Example; Fig 1; 35pp; English.
 XX
 CC The present sequence is encoded by a nitrogen regulatory PII gene. Novel
 CC plant PII (also called P-II) nucleotide sequences have been isolated.
 CC They are useful for regulating nitrogen assimilation in plants, and in
 CC transgenic plant production. They are also used to engineer organisms
 CC that overexpress wild-type or mutant P-II regulatory proteins. P-II
 CC proteins are useful for in vitro screening of herbicides. P-II
 CC nucleotides may be used as probes for isolating additional genomic clones
 CC with the promoters of P-II genes. P-II promoters are light- and/or
 CC sucrose-inducible, and are suitable for genetic engineering of plants
 XX
 SQ Sequence 112 AA;
 Query Match 47.1%; Score 264.5; DB 4; Length 112;
 Best Local Similarity 48.2%; Pred. No. 4.2e-25;
 Matches 53; Conservative 25; Mismatches 31; Indels 1; Gaps 1;
 QY 3 KVEAIVPRWRIQQVSSALLKIGRTVSDVRGFGAQQGSTERHGGSEFSEDKFVAKVKM 62
 DB 3 KVEAIIKPKLDEKVEALQEAQIGLSVIEVKGFGKRGKHTLYRGABYVD-FLPKVKI 61
 QY 63 EIVVKQDVESVINTIIEGARTGEIGDKIFVLPVSDVIRVTRGERGEKA 112
 DB 62 EMVLPDEMVDIAIEAIVGAARTEKIGDKIFVSSIEQAIRITGTGEDA 111

RESULT 10
 ADA34150
 ID ADA34150 standard; protein; 124 AA.
 XX
 AC ADA34150;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Acinetobacter baumannii protein #1311.
 XX

KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
 KW plant biocontrol agent.
 XX
 OS Acinetobacter baumannii.
 XX
 PN US6562958-B1.
 XX
 PD 13-MAY-2003.
 XX
 PF 04-JUN-1999; 99US-00328352.
 XX
 PR 09-JUN-1998; 98US-0088701P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Breton G, Bush D;
 XX
 DR WPI; 2003-576092/54.
 XX
 DR N-PSDB; ADA30024.
 XX
 PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 PT for diagnosing a bacterial disease, as components of antibacterial
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 PT plants.
 XX
 PS Example; SEQ ID NO 5437; 328pp; English.
 XX
 CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents
 CC for diagnosing a bacterial disease, as components of antibacterial
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 CC A. baumannii and other Acinetobacter species in a sample, in screening
 CC compounds for the ability to interfere with the A. baumannii life cycle
 CC or to inhibit A. baumannii infection, and as biocontrol agents for
 CC plants. The present sequence represents the amino acid sequence of an A.
 CC baumannii protein.
 XX
 SQ Sequence 124 AA;
 Query Match 46.4%; Score 260.5; DB 6; Length 124;
 Best Local Similarity 48.6%; Pred. No. 1.5e-24;
 Matches 53; Conservative 23; Mismatches 32; Indels 1; Gaps 1;
 QY 4 VEAIVPRWRIQQVSSALLKIGRTVSDVRGFGAQQGSTERHGGSEFSEDKFVAKVKME 63
 DB 16 VTAIVKPKLDDVREALSDIGVGGITVTEVKGFGKRGKHTLYRGABYVD-FLPKVKIE 74
 QY 64 IIVKQDVESVINTIIEGARTGEIGDKIFVLPVSDVIRVTRGERGEKA 112
 DB 75 LAISDEMVDVAVIESITRVASTGKIGDGKIFVTLNQVIRIRIGETGPD 123

RESULT 11
 AAB69499
 ID AAB69499 standard; protein; 111 AA.
 XX
 AC AAB69499;
 XX
 DT 23-APR-2001 (first entry)
 XX
 DE Rhizobium leguminosarum PII protein.
 XX
 KW Rhizobium leguminosarum; PII; plant nitrogen regulatory gene; P-II;
 KW nitrogen assimilation; transgenic plant; herbicide screening.
 XX
 OS Rhizobium leguminosarum.
 XX
 PN US6177275-B1.
 XX
 PD 23-JAN-2001.
 XX
 PF 23-JUL-1997; 97US-00899330.
 XX

PR 24-JUL-1996; 96US-0022328P.
XX (UJNY) UNIV NEW YORK STATE.
XX Coruzzi GM, Lam H, Hsieh M;
XX WPI; 2001-158572/16.
XX Novel P-II genes capable of regulating plant nitrogen assimilation,
PT useful for transgenic plant production, and as probes for isolating
PT additional genomic clones having P-II gene promoters.
XX Example; Fig 1; 35pp; English.
XX
XX The present sequence is encoded by a nitrogen regulatory PII gene. Novel
CC plant PII (also called P-II) nucleotide sequences have been isolated.
CC They are useful for regulating nitrogen assimilation in plants, and in
CC transgenic plant production. They are also used to engineer organisms
CC that overexpress wild-type or mutant P-II regulatory proteins. P-II
CC proteins are useful for in vitro screening of herbicides. P-II
CC nucleotides may be used as probes for isolating additional genomic clones
CC with the promoters of P-II genes. P-II promoters are light- and/or
CC sucrose-inducible, and are suitable for genetic engineering of plants
XX
XX Sequence 111 AA;
SQ
Query Match 45.8%; Score 257; DB 4; Length 111;
Best Local Similarity 47.7%; Pred. No. 3.7e-24;
Matches 51; Conservative 27; Mismatches 27; Indels 2; Gaps 2;
QY 3 KVEALVRPWRIOQVSSALLKIGIRGVTVSDVRGFGAQQGSTERHGGSEFSDKFAVKVM 62
DB 3 KIEAIKPKLDEVRSF-SGVGLQGIVTEAKGFGKQKGTEDYRGAETVD-FLPKVKV 60
QY 63 EIWVKDQVESVINTIIEGARTGEIGDGKIFVLPVSDVIRVIRGEG 109
DB 61 EVVLADENAEVTEARKAAQTGRIGDGKIFVSNVEVIRVIRGEG 107
RESULT 12
ABG24254
ID ABG24254 standard; protein; 160 AA.
XX AC ABG24254;
XX 18-FEB-2002 (first entry)
XX Novel human diagnostic protein #24245.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-050093631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac R, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX N-PSDB; AAS88441.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensic, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess

PT biodiversity.
XX Claim 20; SEQ ID NO 54613; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have application of mutations
CC diagnostics, forensic, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 160 AA;
SQ
Query Match 45.7%; Score 256.5; DB 4; Length 160;
Best Local Similarity 45.9%; Pred. No. 6.9e-24;
Matches 50; Conservative 28; Mismatches 30; Indels 1; Gaps 1;
QY 4 VEAIVRPWRIOQVSSALLKIGIRGVTVSDVRGFGAQQGSTERHGGSEFSDKFAVKYME 63
DB 52 VTVIIKPFKLDVRSF-SGVGLQGIVTEAKGFGKQKGTEDYRGAETVD-FLPKVKID 110
QY 64 IVVKDQVESVINTIIEGARTGEIGDGKIFVLPVSDVIRVIRGEGKA 112
DB 111 VAIDQDLDEVIDIVSKAAVTGKIGDGKIFVBLQVIRVIRGEGAAA 159
RESULT 13
AAB69500
ID AAB69500 standard; protein; 111 AA.
XX AC AAB69500;
XX 23-APR-2001 (first entry)
XX Bradyrhizobium japonicum PII protein.
XX Bradyrhizobium japonicum; PII; plant nitrogen regulatory gene; P-II;
XX nitrogen assimilation; transgenic plant; herbicide screening.
XX Bradyrhizobium japonicum.
XX US6177275-B1.
XX 23-JAN-2001.
XX 23-JUL-1997; 97US-00899330.
XX 24-JUL-1996; 96US-0022328P.
XX (UJNY) UNIV NEW YORK STATE.
XX Coruzzi GM, Lam H, Hsieh M;
XX WPI; 2001-158572/16.
XX Novel P-II genes capable of regulating plant nitrogen assimilation,
PT useful for transgenic plant production, and as probes for isolating
PT additional genomic clones having P-II gene promoters.

XX Example; Fig 1; 35pp; English.

XX The present sequence is encoded by a nitrogen regulatory PII gene. Novel

XX plant PII (also called P-II) nucleotide sequences have been isolated.

XX They are useful for regulating nitrogen assimilation in plants, and in

XX transgenic plant production. They are also used to engineer organisms

XX that overexpress wild-type or mutant P-II regulatory proteins. P-II

XX proteins are useful for in vitro screening of herbicides. P-II

XX nucleotides may be used as probes for isolating additional genomic clones

XX with the promoters of P-II genes. P-II promoters are light- and/or

XX sucrose-inducible, and are suitable for genetic engineering of plants

SQ Sequence 112 AA;

Query Match 45.6%; Score 256; DB 4; Length 111;

Best Local Similarity 47.7%; Pred. No. 4.9e-24;

Matches 51; Conservative 26; Mismatches 28; Indels 2; Gaps 2;

QY 3 KVEATVPWRIOQVSSALLKIGIRGVTVSDVRGFGAGGSTERHGGSEDFKFAKVM 62

Db 3 KREALIKFKLEDEVS-LSGVGLQGLTVTEAKGFRGKHGHTLYRGAEIVD-FLPKVKI 60

QY 63 EIVKKQVSEVINTIIEGARTGEIGDGKIFVLVPSDVIRVTRG 109

Db 61 EIVIGDGLVERAIDAIRRAAQTRIGDGKIFVSNIEAIRIRTGSG 107

RESULT 14

AAB79181

ID AAB79181 standard; protein; 89 AA.

XX AAB79181;

AC AAB79181;

XX 30-APR-2001 (first entry)

DE Corynebacterium glutamicum HA protein sequence SEQ ID NO:318.

XX Corynebacterium glutamicum; homeostasis; adaptation; HA protein;

KW fine chemical production; organic acid; proteinogenic amino acid;

KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;

KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;

KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;

KW diagnosis; Corynebacterium diphtheriae; genetic engineering;

KW Brevibacterium; environmental condition.

XX Corynebacterium glutamicum.

OS WO200100842-A2.

XX 04-JAN-2001.

XX 23-JUN-2000; 2003WO-IB000911.

XX 25-JUN-1999; 99US-0141031P.

PR 08-JUL-1999; 99CB-01031636.

PR 09-JUL-1999; 99DE-01032125.

PR 09-JUL-1999; 99DE-01032126.

PR 09-JUL-1999; 99DE-01032127.

PR 09-JUL-1999; 99DE-01032128.

PR 09-JUL-1999; 99DE-01032129.

PR 09-JUL-1999; 99DE-01032226.

PR 14-JUL-1999; 99DE-01032920.

PR 14-JUL-1999; 99DE-01032922.

PR 14-JUL-1999; 99DE-01032924.

PR 14-JUL-1999; 99DE-01032928.

PR 14-JUL-1999; 99DE-01032930.

PR 14-JUL-1999; 99DE-01032933.

PR 14-JUL-1999; 99DE-01032935.

PR 14-JUL-1999; 99DE-01032973.

PR 14-JUL-1999; 99DE-01033002.

PR 14-JUL-1999; 99DE-01033003.

PR 14-JUL-1999; 99DE-01033005.

PR 14-JUL-1999; 99DE-01033006.

PR 31-AUG-1999; 99DE-01041378.

PR 31-AUG-1999; 99DE-01041379.

PR 31-AUG-1999; 99DE-01041390.

PR 31-AUG-1999; 99DE-01041391.

PR 03-SEP-1999; 99DE-01042088.

XX (BADI) BASF AG.

XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;

XX WPI; 2001-061974/07.

DR N-PSDB; AAF71296.

XX New isolated Corynebacterium glutamicum nucleic acid for production or

PT modulation of production of fine chemicals such as amino acids, vitamins or

PT nucleosides, nucleotides, lipids, fatty acids, carbohydrates, vitamins or

PT enzymes.

XX Claim 20; Page 563; 712pp; English.

XX AAF71138 to AAF71357 encode the Corynebacterium glutamicum homeostasis

CC and adaptation (HA) proteins given in AAB79023 to AAB79242. The C.

CC glutamicum HA genes (I) can be used in vectors for expression in host

CC cells and production of fine chemicals, such as, an organic acid,

CC proteinogenic or nonproteinogenic amino acid (preferred), purine or

CC pyrimidine base, nucleoside, nucleotide, lipid, saturated or unsaturated

CC fatty acid, diol, carbohydrate, aromatic compound, vitamin, cofactor;

CC polyketide or enzyme. The amino acids produced can be lysine, glutamine,

CC glutamate, alanine, aspartate, glycine, serine, threonine, methionine,

CC cysteine, valine, leucine, isoleucine, arginine, proline, histidine,

CC tyrosine, phenylalanine, or tryptophan. The fine chemical production can

CC be modulated. The presence of (I) or HA proteins encoded by then are used

CC for diagnosing the presence or activity of Corynebacterium diphtheriae.

CC (I) can be used to map the C. glutamicum genome or can be used as markers

CC for genetically engineered Corynebacterium or Brevibacterium. The HA

CC proteins encoded by the (I) are used to maintain homeostasis in C.

CC glutamicum or help the microorganism to adapt to different environmental

CC conditions

XX Sequence 89 AA;

SQ Query Match 43.9%; Score 246.5; DB 4; Length 89;

Best Local Similarity 50.6%; Pred. No. 5.7e-23;

Matches 45; Conservative 24; Mismatches 19; Indels 1; Gaps 1;

QY 24 GIRGVTVSDVRGFGAGGSTERHGGSEDFKFAKVMIEVVKQVSEVINTIIEGAR 83

Db 1 GVQGMVTVTEQGFGQKGHTVYRGAEYAVD-FVPKVKIEVIISDAQAEVINIIVETAR 59

QY 84 TGEIGDGKIFVLVPSDVIRVTRGGEKA 112

Db 60 TGVGVDGKVMNTNIELVVRVTRGGEAA 88

RESULT 15

ABBS4937

ID ABBS4937 standard; protein; 112 AA.

XX ABBS4937;

AC ABBS4937;

XX 29-AUG-2003 (revised)

DT 16-MAY-2002 (first entry)

XX Lactococcus lactis protein glbB.

DE Lactococcus lactis; IL1403.

XX Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.

KW Lactococcus lactis; IL1403.

XX FR2807446-A1.

XX 12-OCT-2001.

```
XX 11-APR-2000; 2000FR-00004630.
PF
XX 11-APR-2000; 2000FR-00004630.
PR
XX (INRG ) INRA INST NAT RECH AGRONOMIQUE.
XX PA
XX Bolotine A, Sorokine A, Renault P, Ehrlich SD;
XX PI
XX WPI; 2002-043418/06.
XX
XX New nucleotide sequence useful in the identification or Lactococcus
PT lactic acid and related species.
XX
XX Claim 6; SEQ ID NO 1639; 2504pp; French.
XX
XX The present invention is related to a Lactococcus lactis nucleotide
CC sequence (ABA90521) and related proteins (ABBS3300-ABBS5621). The nucleic
CC acid sequence is useful in the detection and/or amplification of nucleic
CC acid sequence, particularly to identify Lactococcus lactis or related
CC species. The proteins of the invention are useful for the biosynthesis or
CC biodegradation of a composition of interest. The invention helps research
CC in lactic bacteria, particularly useful in the production of yogurt and
CC cheese. Note: The sequence data for this patent is based on equivalent
CC patent WO200177334 (published 18-OCT-2001) which is available in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences. (Updated on 29-AUG-2003 to
CC standardise OS field)
XX
XX SQ Sequence 112 AA;
Query Match 40.1%; Score 225; DB 5; Length 112;
Best Local Similarity 43.2%; Pred. No. 4e-20;
Matches 48; Conservative 25; Mismatches 36; Indels 2; Gaps 2;
QY 3 KVEALVRPWRHQVSSALLKIG-IRGVTVSDVRGFGAQQGSTERHGGSEFSEDKFVAKVK 61
Db 3 KIEAIITDKLEDLKAALSNDGLVHGVTVSQVLGYGQKGFTYVRGRI-ETTLSSKLK 61
QY 62 MEIVVKDQVESVINTIEGARTGEIGDKIFVLPSDVIKVRVTGERGEKA 112
Db 62 IEIVSIDEKYDDIWNVIKAVQTGEVGDGKIFTQPVVERVIRTSSEDAQA 112
```

Search completed: May 24, 2004, 19:13:02
Job time : 53 secs

Result No.	Query			DB	ID	Description
	Match	Score	Length			
1	100.0	561	196	2	D85024	P II nitrogen sens
2	58.4	327.5	112	2	F70310	nitrogen regulator
3	56.8	318.5	112	2	AH2095	nitrogen regulator
4	55.0	308.5	112	2	H81961	nitrogen regulator
5	54.8	307.5	112	2	B81019	nitrogen regulator
6	54.6	306.5	112	2	A39696	nitrogen regulator
7	51.6	289.5	114	2	D82102	nitrogen regulator
8	51.2	287.5	112	2	S13078	nitrogen regulator
9	50.7	284.5	112	1	RG8CP2	nitrogen regulator
10	50.7	284.5	112	2	S04377	nitrogen regulator
11	50.7	284.5	112	2	C91056	hypothetical prote
12	50.7	284.5	112	2	AF0354	nitrogen regulator
13	50.7	284.5	112	2	AH0826	nitrogen regulator
14	50.4	282.5	112	2	C64307	probable nitrogen
15	50.2	281.5	112	2	G85900	hypothetical prote
16	49.8	279.5	112	2	S73175	nitrogen regulator
17	49.3	276.5	112	2	S53328	nitrogen regulator
18	49.3	276.5	112	2	S33180	nitrogen regulator
19	49.3	276.5	112	2	G64467	nitrogen regulator
20	49.1	275.5	112	2	AB3794	nitrogen regulator
21	49.1	275.5	157	2	A97573	nitrogen regulator
22	48.9	274.5	112	2	F64062	nitrogen regulator
23	48.9	274.5	112	2	AD3374	nitrogen regulator
24	48.8	273.5	112	2	C87493	nitrogen regulator
25	48.6	272.5	112	2	G70747	probable nitrogen
26	48.6	272.5	121	2	B82302	nitrogen regulator
27	46.8	262.5	112	2	T35668	nitrogen regulator
28	46.6	261.5	115	2	D63188	nitrogen regulator
29	46.4	260.5	112	2	B69468	nitrogen regulator

RESULT 4
H81961
nitrogen regulatory protein P-II NMA0447 [similarity] - Neisseria meningitidis (strain Z
C/Spectes: Neisseria meningitidis
C/Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C/Accession: H81961
R/Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A/Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A/Reference number: A81775; MUID:20222556; PMID:10761919
A/Accession: H81961
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-112 <PAR>
A/Cross-references: GB:AL162753; GB:AL157959; NID:g7379120; PIDN:CAB83744.1; PID:g737919
A/Experimental source: serogroup A, strain Z2491
C/Genetics:
A/Gene: glnB; NMA0447
C/Superfamily: regulatory protein P-II
C/Keywords: phosphoprotein
F/51/Binding site: UMP (γ) (covalent) #status predicted

	Query Match	54.8%	Score 307.5;	DB 2;	Length 112;
	Best Local Similarity Matches	54.5%;	Pred.No. 3.2e-23;		
		Conservative 60;	Mismatches 24;	Indels 1;	Gaps 1;
Qy	3 KVRALVPRWRIQQVSSALLKIGIRGVTVSDVRGFGAQQGGSTERHGGSEFGEKFKVAUKKM	62			
Dd	3 KIEALVPFKLLDVREALTEIIGITGMTVSVKVGFRKGHTGYRGAEYAVD-FLPKIKI	61			
Qy	63 EIVKKQKVQESVINTIIEGARTGEIGDGKIFVLPSVDVIRVTGERGEKA	112			
Dd	62 ELVLADDAVERAIIDVIVEARSGIKGDKIFVLPSVEAIRIRTGERSDAA	111			

RESULT 6

A39696

nitrogen regulatory protein P-II - *Synechococcus* sp. (strains PCC 7942 and PCC 6301)

C:Species: *Synechococcus* sp.

A:Variety: PCC 7942; PCC 6301

C:Date: 08-Nov-1991 #sequence_revision 08-Nov-1991 #text_change 26-Aug-1999

C:Accession: A39696 #55817

R:Tsinoernas, N.F.; Casets, A.M.; Harrison, M.A.; Allen, J.P.; Tandeau de Marsac, N.

Proc. Natl. Acad. Sci. U.S.A. 88, 4565-4569, 1991

A:Title: Photosynthetic electron transport controls nitrogen assimilation in cyanobacte

A:Reference number: A39696; PMID:91271233; PMID:1905010

A:Accession: A39696

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-112 <TSI>

A:Cross-references: GB:M62447; NID:G154517; PIDN:AAA27312.1; PID:G552028

A:Experimental source: PCC 7942

A>Note: the authors translated the codon CCT for residue 83 as Thr

R:Li, N.; Warren, P.V.; Golbeck, J.H.; Frank, G.; Zuber, H.; Bryant, D.A.

Biochim. Biophys. Acta 1059, 215-225, 1991

A:Title: Polypeptide composition of the Photosystem I complex and the Photosystem I cor

A:Reference number: A56817; MUID:913555213; PMID:1653017
A:Accession: F56817
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-40 <LI1>
A:Experimental source: PCC 6301
C:Genetics:
A:Gene: glbB
C:Function:
A:Description: involved in signal transduction
A:Note: modification state depends on the nitrogen source and the spectral light quality
C:Superfamily: regulatory protein P-II
C:Keywords: phosphoprotein; signal transduction
E:49/Binding site: Phosphate (Ser) (covalent) #status predicted
F:51/Binding site: UMP (Tyr) (covalent) #status predicted

Query Match 54.6%; Score 306.5; DB 2; Length 112;
Best Local Similarity 55.5%; Pred. No. 4e-23;
Matches 61; Conservative 23; Mismatches 25; Indels 1; Gaps 1;

Qy 3 KVEAIVRPMRIQQVSSALLKIGIRGVTVSDVRGFGAQQGSTERHGGSEFSEDKFVAKVM 62
Db 3 KIEAIRPFKLDVKNALVNAVIGVMTVSEVRGFGKQGTIRYRGSEYTVF-FLQKLL 61
Qy 63 EIVVKDQVESVINTIIEGARTGEIGDGKIFVLPVSDVIRVTRGERGEKA 112
Db 62 EIVVEDAQDVIDKIVAAARPEIGDGKIFVSPDQITIRITGEKNADA 111

RESULT 7
D82102
nitrogen regulatory protein P-II [similarity] - Vibrio cholerae (strain N16961 serogroup
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 24-Aug-2001
C:Accession: D82102
R:Heidelbergl, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
L. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406: 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: D82102
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-114 <HEI>
A:Cross-references: GB:AE004295; GB:AE003852; NID:93656789; PIDN:AAF95383.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Map position: 1
A:Superfamily: regulatory protein P-II
C:Keywords: phosphoprotein
F:53/Binding site: UMP (Tyr) (covalent) #status predicted

Query Match 51.6%; Score 289.5; DB 2; Length 114;
Best Local Similarity 51.8%; Pred. No. 1.9e-21;
Matches 57; Conservative 24; Mismatches 28; Indels 1; Gaps 1;

Qy 3 KVEAIVRPMRIQQVSSALLKIGIRGVTVSDVRGFGAQQGSTERHGGSEFSEDKFVAKVM 62
Db 5 KIEAIRPFKLDVKNALVNAVIGVMTVSEVRGFGKQGTIRYRGSEYTVF-FLPQVKL 63
Qy 63 EIVVKDQVESVINTIIEGARTGEIGDGKIFVLPVSDVIRVTRGERGEKA 112
Db 64 EIVVTDVADRCVDTIIEGARTGEIGDGKIFVLPVSDVIRVTRGERGEKA 113

RESULT 8
SL3078
nitrogen regulatory protein P-II - Azospirillum brasilense
C:Species: Azospirillum brasilense
C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999
C:Accession: SL3078

A:Reference number: A56817; MUID:913555213; PMID:1653017
A:Accession: F56817
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-40 <LI1>
A:Experimental source: PCC 6301
C:Genetics:
A:Gene: glbB
C:Function:
A:Description: involved in signal transduction
A:Note: modification state depends on the nitrogen source and the spectral light quality
C:Superfamily: regulatory protein P-II
C:Keywords: phosphoprotein; signal transduction
E:49/Binding site: Phosphate (Ser) (covalent) #status predicted
F:51/Binding site: UMP (Tyr) (covalent) #status predicted

Query Match 51.2%; Score 287.5; DB 2; Length 112;
Best Local Similarity 49.1%; Pred. No. 3e-21;
Matches 54; Conservative 27; Mismatches 28; Indels 1; Gaps 1;

Qy 3 KVEAIVRPMRIQQVSSALLKIGIRGVTVSDVRGFGAQQGSTERHGGSEFSEDKFVAKVM 62
Db 3 KIEAIRPFKLDVKNALVNAVIGVMTVSEVRGFGKQGTIRYRGSEYTVF-FLPQVKL 61
Qy 63 EIVVKDQVESVINTIIEGARTGEIGDGKIFVLPVSDVIRVTRGERGEKA 112
Db 62 EIVVEDAQDVIDKIVAAARPEIGDGKIFVSPDQITIRITGEKNADA 111

RESULT 9
RGBCP2
nitrogen regulatory protein P-II.1 - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 30-Jun-1988 #sequence_revision 10-Nov-1995 #text_change 01-Mar-2002
C:Accession: C49940; S15991; A29307; S37753; H65032; S31961
R:Liou, J.; Magasanik, B.
J. Bacteriol. 175, 7441-7449, 1993
A:Title: The glbB region of the Escherichia coli chromosome.
A:Reference number: A49940; MUID:94042920; PMID:8226691
A:Accession: C49940
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-112 <LIU>
A:Cross-references: GB:S67014; NID:9455660; PIDN:AAB28779.1; PID:9455663
A:Note: sequence extracted from NCBI backbone (NCBI:139878, NCBIP:139882)
R:Vasudevan, S.G.; Armarego, W.L.F.; Shaw, D.C.; Lilley, P.E.; Dixon, N.E.; Poole, R.K.
Mol. Gen. Genet. 226, 49-58, 1991
A:Title: Isolation and nucleotide sequence of the hmp gene that encodes a haemoglobin-li
A:Reference number: S15991; MUID:91238719; PMID:2034230
A:Accession: S15991
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-112 <VAS>
A:Cross-references: GB:X53872; NID:941730; PIDN:CAA41683.1; PID:941732
A:Experimental source: strain K-12
R:Son, H.S.; Rhee, S.G.
J. Biol. Chem. 262, 8690-8695, 1987
A:Title: Cascade control of Escherichia coli glutamine synthetase: purification and prop
A:Reference number: A29307; MUID:87250488; PMID:2885322
A:Accession: A29307
A:Molecule type: DNA
A:Residues: 1-18, 'R', '20-80, 'E', '83-102, 'VP', <SON>
A:Cross-references: GB:M16778; NID:9146165; PIDN:AAA23883.1; PID:9146166
R:van Heeswijk, W.C.; Rabenberg, M.; Westerhoff, H.V.; Kahn, D.
Mol. Microbiol. 9, 443-457, 1993
A:Title: The genes of the glutamine synthetase adenyllylation cascade are not regulated b
A:Reference number: S36254; MUID:94018640; PMID:8412694
A:Accession: S37753
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-12 <VAN>

A;Cross-references: EMBL:Z21843; NID:G49395; PIDN:CAA79890.1; PID:G49397
 A;Experimental source: strain K-12, substrain W3110
 R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A;Title: The complete genome sequence of *Escherichia coli* K-12.
 A;Reference number: A64720; MUID:97426617; PMID:9278503
 A;Accession: H65032
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-112 <BLAT>
 A;Cross-references: GB:AE000341; GB:U00096; NID:G178899; PIDN:AA75606.1; PID:G1788904;
 A;Experimental source: strain K-12, substrain MGL655
 C;Genetics:
 A;Gene: glnB
 C;Function: <GEN>
 A;Pathway: nitrogen regulation pathway
 A;Note: this protein participates indirectly in activation of glutamine synthase activi
 A;Note: under nitrogen-limiting conditions it is covalently uridylylated; in nitrogen ex
 transferase / uridylyl-removing enzyme which is regulated by the glutamine/alpha-ketogl
 C;Function: <NTRB>
 A;Description: de-uridylylated P-II forms a complex with nitrogen regulation protein II
 the uridylylated form of P-II does not complex with ntrB; free ntrB phosphorylates ntrC
 A;Note: phosphorylated nitrogen regulation protein I (ntrC) activates transcription of t
 C;Function: <NTR>
 A;Description: uridylylated P-II forms a complex with adenyllyltransferase; this complex
 ylyltransferase deadenylates glutamine synthase
 A;Note: free glutamine synthase is active; adenyllylated glutamine synthase is inactive
 C;Superfamily: regulatory protein P-II
 C;Keywords: phosphoprotein; signal transduction
 F;51/Binding site: UMP (Tyr) (covalent) #status experimental
 Query Match 50.7%; Score 284.5; DB 1; Length 112;
 Best Local Similarity 50.9%; Pred. No. 5.9e-21;
 Matches 56; Conservative 27; Mismatches 26; Indels 1; Gaps 1;
 QY 3 KVEAIVRPWRIQQVSSALLKIGIRGVTVSDVRGFGAQQGSTERHGGSEFSDKFKVAKVM 62
 Db 3 KIDAIKPFKLLDDVREALAVGIGTMTVEVKFGGRQKGTLYRGAEYVMD-FLPKVKI 61
 QY 63 EIVVKQDQVESVINTIIEGARTGEIGDKIFVLPVSDVIRVIRGTGERGEKA 112
 Db 62 EIVVPDDIVDTCVDTIIRTATQTKIGDKIFVDFVARVIRIRTGEDDAA 111
 RESULT 10
 S04377
 nitrogen regulatory protein P-II - *Klebsiella pneumoniae*
 C;Species: *Klebsiella pneumoniae*
 C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 26-Aug-1999
 C;Accession: S04377
 R;Holtel, A.; Merrick, M.
 Mol. Gen. Genet. 215, 134-138, 1988
 A;Title: Identification of the *Klebsiella pneumoniae* glnB gene: nucleotide sequence of w
 A;Reference number: S04376; MUID:89201233; PMID:2907369
 A;Accession: S04377
 A;Molecule type: DNA
 A;Residues: 1-112 <HOL>
 A;Cross-references: EMBL:X14312; NID:G43804; PIDN:CAA32177.1; PID:G43806
 A;Experimental source: strain M5al
 C;Genetics:
 A;Gene: glnB
 C;Function:
 A;Description: important for the control of glutamine synthase; in nitrogen-limiting con
 to form P-II-UMP; P-II-UMP allows the deadenylation of glutamine synthetase, thus activa
 GS
 A;Note: P-II indirectly controls the transcription of the GS gene (glnA) by preventing N
 ted, these events are reversed
 C;Superfamily: regulatory protein P-II
 C;Keywords: phosphoprotein; signal transduction
 F;51/Binding site: UMP (Tyr) (covalent) #status predicted
 Query Match 50.7%; Score 284.5; DB 2; Length 112;

Best Local Similarity 50.9%; Pred. No. 5.9e-21;
 Matches 56; Conservative 27; Mismatches 26; Indels 1; Gaps 1;
 QY 3 KVEAIVRPWRIQQVSSALLKIGIRGVTVSDVRGFGAQQGSTERHGGSEFSDKFKVAKVM 62
 Db 3 KIDAIKPFKLLDDVREALAVGIGTMTVEVKFGGRQKGTLYRGAEYVMD-FLPKVKI 61
 QY 63 EIVVKQDQVESVINTIIEGARTGEIGDKIFVLPVSDVIRVIRGTGERGEKA 112
 Db 62 EIVVTDDIVDTCVDTIIRTATQTKIGDKIFVDFVARVIRIRTGEDDAA 111
 RESULT 11
 C91056
 hypothetical protein ECs3419 [imported] - *Escherichia coli* (strain O157:H7, substrain R;
 C;Species: *Escherichia coli*
 C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
 C;Accession: C91056
 R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A;Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and gen
 A;Reference number: A99629; MUID:21156231; PMID:11258796
 A;Accession: C91056
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-112 <HAY>
 A;Cross-references: GB:BA000007; PIDN:BA36842.1; PID:G13362890; GSPDB:GN00154
 A;Experimental source: strain O157:H7, substrain RIMD 0509952
 C;Genetics:
 A;Gene: ECs3419
 C;Superfamily: regulatory protein P-II
 Query Match 50.7%; Score 284.5; DB 2; Length 112;
 Best Local Similarity 50.9%; Pred. No. 5.9e-21;
 Matches 56; Conservative 27; Mismatches 26; Indels 1; Gaps 1;
 QY 3 KVEAIVRPWRIQQVSSALLKIGIRGVTVSDVRGFGAQQGSTERHGGSEFSDKFKVAKVM 62
 Db 3 KIDAIKPFKLLDDVREALAVGIGTMTVEVKFGGRQKGTLYRGAEYVMD-FLPKVKI 61
 QY 63 EIVVKQDQVESVINTIIEGARTGEIGDKIFVLPVSDVIRVIRGTGERGEKA 112
 Db 62 EIVVPDDIVDTCVDTIIRTATQTKIGDKIFVDFVARVIRIRTGEDDAA 111
 RESULT 12
 AF0354
 nitrogen regulatory protein P-II glnB [imported] - *Yersinia pestis* (strain CO92)
 C;Species: *Yersinia pestis*
 C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
 C;Accession: AF0354
 R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 ll, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,
 Nature 413, 523-527, 2001
 A;Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
 A;Reference number: AB0001; MUID:21470413; PMID:11586360
 A;Accession: AF0354
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-112 <KUR>
 A;Cross-references: GB:AL590842; PIDN:CAC92161.1; PID:G15980875; GSPDB:GN00175
 C;Genetics:
 A;Gene: glnB
 C;Superfamily: regulatory protein P-II
 Query Match 50.7%; Score 284.5; DB 2; Length 112;
 Best Local Similarity 50.0%; Pred. No. 5.9e-21;
 Matches 55; Conservative 28; Mismatches 26; Indels 1; Gaps 1;
 QY 3 KVEAIVRPWRIQQVSSALLKIGIRGVTVSDVRGFGAQQGSTERHGGSEFSDKFKVAKVM 62
 Db 3 KIDAIKPFKLLDDVREALAVGIGTMTVEVKFGGRQKGTLYRGAEYVMD-FLPKVKI 61

Db 3 KVEAIIIRPKLFIIVKKALSDAGYGMVTSEVKGQGVGGGIVEGRGYVD-LIPKVKI 61

Qy 93 EIWVKKDQVESINTIECARTGETGDGI FVL PYSDVIRVRTGBRGEKA 112
 |::|| :::|| |::|| |::|| |::|| |::||
Db 62 ELVVKEEDVNVIDIICENARIGNFDGDKIFVIPVRVVRFKEEGHEA 111
 |::|| :::|| |::|| |::|| |::|| |::||

RESULT 15

G85900
hypothetical protein glnB [imported] - Escherichia coli (strain O157:H7, subspecies: Escherichia coli)
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: G85900
R.;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.; Miller, L.; Grothbeck, E.J.; Davis, N.W.; Lim, A.; Diallanta, E.; Potamouasis, K. Nature 409, 529-533, 2001.
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: AB5480; PMID:21074935; PMID:11206551
A;Accession: G85900
A>Status: preliminary
A:Molecule type: DNA
A;Residues: 1-112 <STO>
A;Cross-references: GB:AE005174; NID-gt2516963; PIDN:AAC57667.1; GSFPDB:GN001418
A;Experimental source: strain O157:H7, substrain EDU933
C;Genetics:
A;Gene: glnB
C;Superfamily: regulatory protein p-II

[illegible]

Search completed: May 24, 2004, 19:15:40

RESULT 14

Query Match	50.4%	Score	282.5	DB 2	Length	112	
Best Local Similarity	53.6%	Pred. No.	9.2e-21				
Matches	59	Conservative	21	Mismatches	29	Indels	1
				Gaps	1		1

QY 3 KVEAIVRPWRIQOVSSALLKIGIRGTVSDVRGFGAQQGSTERHGGSEFSEDEKFAVKVM 62

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 19:06:17 ; Search time 10.5 Seconds
(without alignments)
560.374 Million cell updates/sec

Title: US-09-756-541-1
Perfect score: 561
Sequence: 1 FYKVEAIRPWRIOQVSSAL.....VLPVSDVIRVTGEGKAE 113

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 3

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	327.5	58.4	112	1 GLNB_AQUAE	O66513 aquifex aeo
2	317.5	56.6	112	1 GLNB_PREDI	Q47894 fremvella d
3	316.5	56.4	112	1 GLNB_NOSPU	O30794 nostoc punc
4	312.5	55.7	112	1 GLNB_SYN7	P80016 synechococc
5	312.5	55.7	112	1 GLNB_SYN3	O55247 synechocyst
6	290.5	51.8	112	1 GLNB_PASMU	Q9CJ41 pasteurella
7	287.5	51.2	112	1 GLNB_AZOB	P21193 azospirillum
8	284.5	50.7	112	1 GLNB_ECOLI	P35826 escherichia
9	284.5	50.7	112	1 GLNB_KLEOX	P11671 klebsiella
10	282.5	50.4	112	1 Y059_METJA	Q60381 methanococc
11	279.5	49.8	112	1 GLNB_PORPU	P51254 porphyra pu
12	276.5	49.3	112	1 GLNB_RHORI	Q53044 rhodospirill
13	276.5	49.3	112	1 GLNB_RHOSH	P43519 rhodobacter
14	276.5	49.3	112	1 YD44_METJA	Q58740 methanococc
15	275.5	49.1	112	1 GLNB_RHIME	O52905 rhizobium m
16	274.5	48.9	112	1 GLNB_HAEIN	P43795 haemophilus
17	272.5	48.6	112	1 GLNB_MYCTU	Q10960 mycobacteri
18	272.5	48.6	112	1 GLNB_RHILLO	Q98n18 rhizobium l
19	268.5	47.9	112	1 GLNB_BRAJA	P14179 bradyrhizob
20	264.5	47.1	112	1 GLNB_RHOCA	P13556 rhodobacter
21	261.5	46.6	112	1 GLN2_METTH	O26760 methanobact
22	257	45.8	111	1 GLNB_RHILV	P09827 rhizobium l
23	256.5	45.7	112	1 GLNK_ECOLI	P38504 escherichia
24	256.5	45.7	113	1 GLNB_CYACA	Q9Em37 cyanidium c
25	251.5	44.8	115	1 GLN1_METTH	O26758 methanobact
26	235.5	42.0	112	1 GLNB_RHIEH	O54053 rhizobium e
27	201.5	35.9	116	1 NRGB_BACSU	Q07428 bacillus su
28	189.5	33.8	123	1 GLN2_METBA	P54809 methanosarc
29	171.5	30.6	128	1 GLN2_METTL	P25770 methanococc
30	169.5	30.2	125	1 GLN4_METBA	P54806 methanosarc
31	169	30.1	122	1 GLN2_METIV	P51604 methanobact
32	167	29.8	121	1 GLN2_METMP	P71525 methanococc
33	161	28.7	121	1 GLN2_METTM	Q50787 methanobact

RESULT 1

ID	GLNB_AQUAE	STANDARD;	PRT;	112 AA.
AC	O66513;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Nitrogen regulatory protein P-II.			
GN	GLNB OR AQ_109.			
OS	Aquifex aeolicus.			
OC	Bacteria; Aquificae;			
OX	NCBI_TaxID=63363;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=VP5;			
RX	MEDLINE=98196666; PubMed=9537320;			
RA	Decker G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L., Graham D.E., Overbeek R., Snead M.A., Keller M., Anjay M., Huber R., Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;			
RA	"The complete genome of the hyperthermophilic bacterium Aquifex aeolicus."			
RT	Nature 392:353-358 (1998).			
RL	FUNCTION: IN NITROGEN-LIMITING CONDITIONS, WHEN THE RATIO OF GLN TO 2-KETOGLUTARATE DECREASES, P-II IS URIDYLATED TO P-II-UMP. P-II-UMP ALLOWS THE DEADENYLATION OF GLUTAMINE SYNTHETASE (GS). THIS ACTIVATING THE ENZYME. CONVERSELY, IN NITROGEN EXCESS P-II IS DEURIDYLATED AND PROMOTES THE ADENYLATION OF GS (BY SIMILARITY).			
CC	FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE GS GENE (GLNA). P-II PREVENTS NR-II CATALYZED CONVERSION OF NR-I TO NR-I-PHOSPHATE. THE TRANSCRIPTIONAL ACTIVATOR OF GLNA. WHEN P-II IS URIDYLATED TO P-II-UMP, THESE EVENTS ARE REVERSED (BY SIMILARITY).			
CC	!- SUBUNIT: Homotrimer (By similarity).			
CC	!- SIMILARITY: Belongs to the P(II) protein family.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).			
CC	EMBL; AE000674; AAC06473.1; -			
DR	PIR; F70310; F70310.			
DR	HSSP; P05826; 2PII.			
DR	InterPro; IPR002187; PII_glnB.			
DR	InterPro; IPR002332; PII_glnB_ump_s.			
DR	Pfam; PF00543; P-II; 1			
DR	PRINTS; PR003440; PIIglnB.			
DR	ProDom; PD001194; PII_glnB.			
DR	PROSITE; PS00496; PII_glnB_ump; FALSE_NEG.			
DR	PROSITE; PS00638; PII_glnB_CTER; 1.			
KW	Transcription regulation; Nitrogen fixation; Complete proteome.			
ET	BINDING 51 51 UMP (BY SIMILARITY).			

P25771 methanococc
P54808 methanosarc
Q50786 methanobact
P51603 methanobact
P71524 methanococc
P54807 methanosarc
P38081 saccharomyc
Q98n24 halobacteri
Q8tvm0 methanopyru
P58865 methanosarc
P05517 bacillus th
P59179 bifidobacte

ALIGNMENTS

SQ SEQUENCE 112 AA; 12497 MW; OE44B4B371A62233B CRC64;
 Query Match 58.4%; Score 327.5; DB 1; Length 112;
 Best Local Similarity 54.5%; Pred. No. 2.5e-26;
 Matches 60; Conservative 30; Mismatches 19; Indels 1; Gaps 1;
 QY 3 KVEAIVRPWRIQOVSSALLKIGIRGVTVSDVRGFGAGGSGTERRHGGSEFSEDFKFAKVM 62
 DB 3 KIEALIKPKLDEKDALVEIGIGMTVTGKVGQKGKGTETVYRGTEYVID-FLPKVKI 61
 QY 63 EIVVKKDQVESINTIIEGARTGEIGDKIFVLVPSDVIRVTRGERGEKA 112
 DB 62 EIVVEDNQDMVWDKIIAAARTGEIGDKIFISPVQVVRIRTKNTEA 111
 QY 63 EIVVKKDQVESINTIIEGARTGEIGDKIFVLVPSDVIRVTRGERGEKA 112
 DB 62 EIVVEDNQDMVWDKIIAAARTGEIGDKIFISPVQVVRIRTKNTEA 111
 RESULT 2
 GLNB_PREDI STANDARD; PRT; 112 AA.
 ID GLNB_FREDI
 AC Q47894;
 DT 01-NOV-1997 (Rel. 35, Created);
 DT 01-NOV-1997 (Rel. 35, Last sequence update);
 DT 01-NOV-1997 (Rel. 35, Last annotation update);
 DE Nitrogen regulatory protein P-II (PII signal transducing protein).
 GN GLNB.
 OS Fremyella diplosiphon (Calothrix PCC 7631).
 OC Bacteria; Cyanobacteria; Nostocales; Microchaetaceae; Fremyella.
 OX NCBI_TaxID=1197;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Liotenberg S., Carsets A.M., Campbell D., Hounard J.,
 RA Tandeau de Marsac N.;
 RA Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE GS
 CC GENE (GLNA). P-II PREVENTS NR-II CATALYZED CONVERSION OF NR-I TO
 CC NR-I-PHOSPHATE. THESE EVENTS ARE REVERSED (BY SIMILARITY).
 CC IS PHOSPHORYLATED, THE TRANSCRIPTIONAL ACTIVATOR OF GLNA. WHEN P-II
 CC TO 2-KETOGLUTARATE DECREASES, P-II IS PHOSPHORYLATED WHICH ALLOWS
 CC THE DEADENYLATION OF GLUTAMINE SYNTHETASE (GS), THUS ACTIVATING
 CC -!- SUBUNIT: Homotrimer (By similarity).
 CC -!- PTM: PHOSPHORYLATION DEPENDENT ON THE NITROGEN SOURCE AND SPECTRAL
 CC LIGHT QUALITY (BY SIMILARITY).
 CC -!- SIMILARITY: Belongs to the P(II) protein family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; X97327; CAA65992.1; --
 DR HSSP; P05826; 2P11.
 DR InterPro; IPR002187; PII_glnB.
 DR InterPro; IPR002332; PII_GlnB_UMP_S.
 DR Pfam; PF00543; P-II; 1.
 DR PRINTS; PR00340; PIIGLNB.
 DR ProDom; PD001194; PII_glnB; 1.
 DR PROSITE; PS00496; PII_GlnB_UMP; 1.
 DR PROSITE; PS00638; PII_GlnB_CTER; 1.
 KW Transcription regulation; Nitrogen fixation; Phosphorylation.
 FT MOD RES 49 49 PHOSPHORYLATION (PROBABLE).
 SQ SEQUENCE 112 AA; 12478 MW; F74543C393C4596B CRC64;
 Query Match 56.6%; Score 317.5; DB 1; Length 112;
 Best Local Similarity 54.4%; Pred. No. 2.6e-25;
 Matches 62; Conservative 25; Mismatches 22; Indels 1; Gaps 1;
 QY 3 KVEAIVRPWRIQOVSSALLKIGIRGVTVSDVRGFGAGGSGTERRHGGSEFSEDFKFAKVM 62
 DB 3 KVEAIVRPWRIQOVSSALLKIGIRGVTVSDVRGFGAGGSGTERRHGGSEFSEDFKFAKVM 62
 QY 63 EIVVKKDQVESINTIIEGARTGEIGDKIFVLVPSDVIRVTRGERGEKA 112
 DB 62 EIVVEDNQDMVWDKIIAAARTGEIGDKIFISPVQVVRIRTKNTEA 111
 RESULT 3
 GLNB_NOSPU STANDARD; PRT; 112 AA.
 ID GLNB_NOSPU
 AC Q30734;
 DT 15-DEC-1998 (Rel. 37, Created);
 DT 15-DEC-1998 (Rel. 37, Last sequence update);
 DT 15-JUL-1999 (Rel. 38, Last annotation update);
 DE Nitrogen regulatory protein P-II (PII signal transducing protein).
 GN GLNB.
 OS Nostoc punctiforme.
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 OX NCBI_TaxID=63737;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=ATCC 29133 / PCC 73102;
 RC MEDLINE=98304077; PubMed=9639924;
 RA Hanson T.B., Forchhammer K., Tandeau de Marsac N., Meeks J.C.;
 RA "Characterization of the glnB gene product of Nostoc punctiforme
 RT strain ATCC 29133: glnB or the PII protein may be essential.";
 RL Microbiology 144:1537-1547(1998).
 CC -!- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE GS
 CC GENE (GLNA). P-II PREVENTS NR-II CATALYZED CONVERSION OF NR-I TO
 CC NR-I-PHOSPHATE. THESE EVENTS ARE REVERSED (BY SIMILARITY).
 CC IS PHOSPHORYLATED, THE TRANSCRIPTIONAL ACTIVATOR OF GLNA. WHEN P-II
 CC TO 2-KETOGLUTARATE DECREASES, P-II IS PHOSPHORYLATED WHICH ALLOWS
 CC THE DEADENYLATION OF GLUTAMINE SYNTHETASE (GS), THUS ACTIVATING
 CC THE ENZYME (BY SIMILARITY).
 CC -!- SUBUNIT: Homotrimer (By similarity).
 CC -!- PTM: PHOSPHORYLATION DEPENDENT ON THE NITROGEN SOURCE AND SPECTRAL
 CC LIGHT QUALITY (BY SIMILARITY).
 CC -!- SIMILARITY: Belongs to the P(II) protein family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; AF017419; AAC26348.1; --
 DR HSSP; P05826; 2P11.
 DR InterPro; IPR002187; PII_glnB.
 DR InterPro; IPR002332; PII_GlnB_UMP_S.
 DR Pfam; PF00543; P-II; 1.
 DR PRINTS; PR00340; PIIGLNB.
 DR ProDom; PD001194; PII_glnB; 1.
 DR PROSITE; PS00496; PII_GlnB_UMP; 1.
 DR PROSITE; PS00638; PII_GlnB_CTER; 1.
 KW Transcription regulation; Nitrogen fixation; Phosphorylation.
 FT MOD RES 49 49 PHOSPHORYLATION (PROBABLE).
 SQ SEQUENCE 112 AA; 12478 MW; 9C2224C38B67583A CRC64;
 Query Match 56.4%; Score 316.5; DB 1; Length 112;
 Best Local Similarity 57.3%; Pred. No. 3.2e-25;
 Matches 63; Conservative 23; Mismatches 23; Indels 1; Gaps 1;
 QY 3 KVEAIVRPWRIQOVSSALLKIGIRGVTVSDVRGFGAGGSGTERRHGGSEFSEDFKFAKVM 62
 DB 3 KVEAIVRPWRIQOVSSALLKIGIRGVTVSDVRGFGAGGSGTERRHGGSEFSEDFKFAKVM 62
 QY 63 EIVVKKDQVESINTIIEGARTGEIGDKIFVLVPSDVIRVTRGERGEKA 112
 DB 62 EIVVEDNQDMVWDKIIAAARTGEIGDKIFISPVQVVRIRTKNTEA 111

CC -! SUBUNIT: Homotrimer.
CC -! PTM: PHOSPHORYLATION DEPENDENT ON THE NITROGEN SOURCE AND SPECTRAL

```

CC EVENTS ARE REVERSED. WHEN THE RATIO OF GLN TO 2-KETOGLUTARATE
CC DECREASES, P-II IS URIDYLATED TO P-II-UMP, WHICH CAUSES THE
CC DEADENYLATION OF GUANININE SYNTHETASE BY GLNE, SO ACTIVATING THE
CC ENZYME (BY SIMILARITY).
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- PTM: Uridylated/deuridylated by glnd (By similarity).
CC -1- SIMILARITY: Belongs to the P(II) protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; AE006236; AAK04088.1; -.
CC DR HSSP; P05826; 2PII.
CC DR InterPro; IPR002187; PII_glnB.
CC DR InterPro; IPR002332; PII_GlnB_UMP_S.
CC DR Pfam; PF00543; P-II; 1.
CC DR PRINTS; PR00340; PIIGLNB.
CC DR ProDom; PDO01194; PII_glnB; 1.
CC DR PROSITE; PS00496; PII_GlnB_UMP; 1.
CC DR PROSITE; PS00638; PII_GlnB_CTER; 1.
CC KW Transcription regulation; Nitrogen fixation; Complete proteome.
CC FT BINDING 51 51 UMP (BY SIMILARITY).
CC SQ SEQUENCE 112 AA; 12684 MW; 0CC9D24651F47642 CRC64;
CC -----
Query Match 51.8%; Score 290.5; DB 1; Length 112;
Best Local Similarity 53.6%; Pred. No. 1.4e-22;
Matches 59; Conservative 21; Mismatches 29; Indels 1; Gaps 1;

QY 3 KVEATVPWRIQOVSSALLKIGIRGTVSDVDFGFGAGGSTERHGSGERSEDFVAKVM 62
Db 3 KIEALIKPFLDDVRESLDWGTMTVTVEVRGFGKQGHTELYRGAEMVD-FLPKVM 61
QY 63 EIVKKXQVQVESVINTIIEGARTGEIGDGKIFVLVPDVIKVRVTRGEGKA 112
Db 62 EIVVTDEQVQDCEIAIMETAQTGKIGDGKIFVYDVVERVIRIRTEENEDA 111

RESULT 7
GlnB_AZOBR
ID GlnB_AZOBR STANDARD; PRT; 112 AA.
AC P21193;
DT 01-MAY-1991 (rel 18 Created)

```

01-NOV-1995 (Rel. 32, Last annotation update)
DE Nitrogen regulatory protein pN-II.
GN ALNB.
OS Azospirillum brasilense.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Rhodospirillaceae; Azospirillum.
NCBI_TaxID=192;
[1] _____
RN SEQUENCE FROM N.A.
RP STRAIN=Sp7;
RC MEDLINE=91094780; PubMed=1702507;
RX de Zamaroczy M., Delorme F., Elmerich C.;
RA "Characterization of three different nitrogen-regulated promoter
RT regions for the expression of glbB and glbA in Azospirillum
RT brasilense.";
RL Mol. Gen. Genet. 224:421-430(1990).
RN [2] _____
RP SEQUENCE OF 88-112 FROM N.A.
RC STRAIN=Sp7;
RX MEDLINE=87076765; PubMed=28786895;
RA Bozouklian H., Elmerich C.;
RT "Nucleotide sequence of the Azospirillum brasilense Sp7 glutamine
RT synthetase structural gene.";
RL Biochimie 68:1181-1187(1986).
RN -!- FUNCTION: IN NITROGEN-LIMITING CONDITIONS. WHEN THE RATIO OF GLN
CC

CC TO 2-KETOGLUTARATE DECREASES, P-II IS URIDYLATED TO P-II-UMP.
CC P-II-UMP ALLOWS THE DEADENYLATION OF GLUTAMINE SYNTHETASE (GS),
CC THUS ACTIVATING THE ENZYME. CONVERSELY, IN NITROGEN EXCESS P-II
CC IS DEURIDYLATED AND PROMOTES THE ADENYLATION OF GS.
CC FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE GS
CC GENE (GLN). P-II PREVENTS NR-II CATALYZED CONVERSION OF NR-I TO
CC NR-I-PHOSPHATE, THE TRANSCRIPTIONAL ACTIVATOR OF GLN. WHEN P-II
CC IS URIDYLATED TO P-II-UMP, THESE EVENTS ARE REVERSED.
CC SUBUNIT: Homotrimer (By similarity).
CC SIMILARITY: Belongs to the P(II) protein family.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X51499; CAA35867.1; -;
CC EMBL; M26107; -; NOT_ANNOTATED_CDS.
CC PIR; S13078; S13078.
CC HSSP; P05826; 2P11.
CC InterPro; IPR002187; PII_glnB.
CC InterPro; IPR002332; PII_glnB_UMP_S.
CC Pfam; PF00543; P-II; 1.
CC PRINTS; PR00340; PIIGLNB.
CC ProDom; PD001194; PII_glnB.
CC PROSITE; PS00496; PII_GLNBL_UMP; 1.
CC PROSITE; PS00638; PII_GLNBL_CTER; 1.
CC Transcription regulation; Nitrogen fixation.
CC BINDING 51 51 UMP.
CC SEQUENCE 112 AA; 12371 MW; B6258EC9C4B63871 CRC64;

Query Match 51.2%; Score 287.5; DB 1; Length 112;
Best Local Similarity 49.1%; Pred. No. 2.8e-22;
Matches 54; Conservative 27; Mismatches 28; Indels 1; Gaps 1;

QY 3 KVEALVPRWICQVSSALLKIGIRGVTVSDVDFGAGQGGSTRHGSGFSEDEKFAVKVM 62
Db 3 KLEAIKPKLDEVEALHEVGKIGITVTEAKGFGKQGHTELRGAETVVD-FLPKVKI 61

QY 63 EIVKKKQDVESVINTIEGARTGIGDGKIFVLPSDVIRVTRGERKA 112
Db 62 EVMWEDSLVERAIEALQQAHTGRIGDGKIFVTEVEVVRIRNGGSDA 111

RESULT 8
GLNB_ECOLI STANDARD; PRT; 112 AA.
AC P05826;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Nitrogen regulatory protein P-II 1.
GN GLNB OR B2553 OR C3076 OR Z3829 OR ECS3419 OR STM2561 OR STY2808 OR
GN T0295 OR SF2600 OR S2772.
OS Escherichia coli.
OS Escherichia coli C6.
OS Escherichia coli C157:H7,
OS Salmonella typhimurium,
OS Salmonella typhi, and
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 217992, 83334, 602, 601, 623;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli;
RX MEDLINE=87250488; PubMed=2885322;
RA Son H.S., Rhee S.G.;
RT Cascade control of Escherichia coli glutamine synthetase.
RT Purification and properties of PII protein and nucleotide sequence of

RT its structural gene.";
RL J. Biol. Chem. 262:8690-8695 (1987).
RN [2]
RP REVISIONS TO C-TERMINUS.
RC SPECIES=E.coli;
RX MEDLINE=89201233; PubMed=2907369;
RA Holtel A., Merrick M.;
RT Identification of the Klebsiella pneumoniae glnB gene: nucleotide
RT sequence of wild-type and mutant alleles.";
RL Mol. Gen. Genet. 215:134-138 (1988).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=KL2;
RX MEDLINE=91238719; PubMed=2034230;
RA Vasudevan S.G., Armarego W.L.F., Shaw D.C., Lilley P.E., Dixon N.E.,
RA Poole R.K.;
RT Isolation and nucleotide sequence of the bmp gene that encodes a
RT haemoglobin-like protein in Escherichia coli K-12.";
RL Mol. Gen. Genet. 226:49-58 (1991).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli;
RX MEDLINE=94042920; PubMed=8226691;
RA Liu J., Magasanik B.;
RT The glnB region of the Escherichia coli chromosome.";
RL J. Bacteriol. 175:7441-7449 (1993).
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=KL2 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474 (1997).
RN [6]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=KL2;
RX MEDLINE=97349980; PubMed=9205837;
RA Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
RA Oshiba T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasubram S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horiuchi T.;
RT Construction of a contiguous 874-kb sequence of the Escherichia coli
RT K-12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features.";
RL DNA Res. 4:91-113 (1997).
RN [7]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=C6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
RN [8]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=0157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai J., Hackett J., Klink S., Boutin A., Shao Y., Miller J.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamasis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533 (2001).

RN SEQUENCE FROM N.A.
 RP SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;
 RC MEDLINE=21156231; PubMed=11258796;
 RX Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 RA Kihara S., Shiba T., Hattori M., Shinagawa H.;
 RA "Complete genome sequence of enterohemorrhagic *Escherichia coli*
 RT O157:H7 and genomic comparison with a laboratory strain K-12";
 RL DNA Res. 8:11-22(2001).
 RN [10]
 RN SEQUENCE OF 1-12 FROM N.A.
 RP SPECIES=E.coli; STRAIN=K12 / W31110;
 RC MEDLINE=94018640; PubMed=8412694;
 RX van Heeswijk W.C., Rabenberg M., Westerhoff H.V., Kahn D.;
 RA "The genes of the glutamine synthetase adenylation cascade are not
 RT regulated by nitrogen in *Escherichia coli*";
 RL Mol. Microbiol. 9:443-458(1993).
 RN [11]
 RN SEQUENCE FROM N.A.
 RP SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RC MEDLINE=21534948; PubMed=11677609;
 RX McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Fiorea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RA "Complete genome sequence of *Salmonella enterica* serovar Typhimurium
 RT LT2";
 RL Nature 413:852-856(2001).
 RN [12]
 RN SEQUENCE FROM N.A.
 RP SPECIES=S.typhi; STRAIN=CT18;
 RC MEDLINE=21534947; PubMed=11677608;
 RX Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagels K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
 RA Quail M.A., Rutherford K., Stevens M., Skelton J., Stevens K.,
 RA Whitehead S., Barrell B.G.;
 RA "Complete genome sequence of a multiple drug resistant *Salmonella*
 RT enterica serovar Typhi CT18";
 RL Nature 413:848-852(2001).
 RN [13]
 RN SEQUENCE FROM N.A.
 RP SPECIES=S.typhi; STRAIN=Ty2 / ATCC 700931;
 RC MEDLINE=22531367; PubMed=12644504;
 RX Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 RA Burland V., Kodyanni V., Schwartz D.C., Blattner F.R.;
 RA "Comparative genomics of *Salmonella enterica* serovar Typhi strains Ty2
 RT and CT18";
 RL J. Bacteriol. 185:2330-2337(2003).
 RN [14]
 RN SEQUENCE FROM N.A.
 RP SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
 RC MEDLINE=22272406; PubMed=12384590;
 RX Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
 RA Yu J.;
 RA "Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity
 RT through comparison with genomes of *Escherichia coli* K12 and O157";
 RL Nucleic Acids Res. 30:4432-4441(2002).
 RN [15]
 RN SEQUENCE FROM N.A.
 RP SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;
 RC MEDLINE=22530274; PubMed=12704152;
 RX Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
 RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,

RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
 RT Schwartz D.C., Blattner F.R.;
 RA "Complete genome sequence and comparative genomics of *Shigella*
 RT flexneri serotype 2a strain 2457T";
 RL Infect. Immun. 71:2775-2786(2003).
 RN [16]
 RN REVIEW.
 RC SPECIES=E.coli;
 RX MEDLINE=90089474; PubMed=2574599;
 RA Magasanik B.;
 RT "Regulation of transcription of the *glnALG* operon of *Escherichia coli*
 RT by protein phosphorylation";
 RL Biochimie 71:1005-1012(1989).
 RN [17]
 RN CRYSTALLIZATION, AND SUBUNITS.
 RC SPECIES=E.coli;
 RX MEDLINE=94123764; PubMed=8293810;
 RA Vasudevan S.G., Gedye C., Dixon N.E., Cheah E., Carr P.D.,
 RA Suffolk P.M., Jeffrey P.D., Ollis D.L.;
 RT "Escherichia coli PII protein: purification, crystallization and
 RT oligomeric structure";
 RL ZEBs Lett. 337:255-258(1994).
 RN [18]
 RN X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
 RC SPECIES=E.coli;
 RX MEDLINE=95171116; PubMed=7866749;
 RA Cheah E., Carr P.D., Suffolk P.M., Vasudevan S.G., Dixon N.E.,
 RA Ollis D.L.;
 RT "Structure of the *Escherichia coli* signal transducing protein PII";
 RL Structure 2:981-990(1994).
 Query Match 50.7%; Score 284.5; DB 1; Length 112;
 Best Local Similarity 50.9%; Pred. No. 5.6e-22;
 Matches 56; Conservative 27; Mismatches 26; Indels 1; Gaps 1;
 Qy 3 KVEAIVPWRIQVSSALLKIGIVTSVDSVGRGCAQGGSTERHGGSFSDKFKVAKYM 62
 Db 3 KIDAIKPKKLDVREALAELVGITGMTVEVKFGKQKGTLYRGAEYWD-FLPKVKI 61
 Qy 63 EIVVKKQVESVINTIIEGARTGEIGDKTFLVPSDVIRVTGERGKA 112
 Db 62 EIVVDDILVDTCDVTIIRTATQTKIGDKGFVDFVARVIRITGEEDDAA 111
 RESULT 9
 GLNB_KLROX STANDARD; PRT; 112 AA.
 ID GLNB_KLROX STANDARD; PRT; 112 AA.
 AC P11671;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Nitrogen regulatory protein P-II.
 GN GLNB.
 OS *Klebsiella oxytoca*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; *Klebsiella*.
 ON NCBI_TaxID=571;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSa1;
 RX MEDLINE=89201233; PubMed=2907369;
 RA Holtel A., Merrick M.;
 RT "Identification of the *Klebsiella pneumoniae* *glnB* gene: nucleotide
 RT sequence of wild-type and mutant alleles";
 RL Mol. Gen. Genet. 215:134-138(1988).
 CC -!- FUNCTION: IN NITROGEN-LIMITING CONDITIONS, WHEN THE RATIO OF GLN
 CC TO 2-KETOGLUTARATE DECREASES, P-II IS URIDYLATED TO P-II-UMP.
 CC P-II-UMP ALLOWS THE DEADENYLATION OF GLUTAMINE SYNTHETASE (GS),
 CC THUS ACTIVATING THE ENZYME. CONVERSELY, IN NITROGEN EXCESS P-II
 CC IS DEURIDYLATED AND PROMOTES THE ADENYLATION OF GS.
 CC -!- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE GS
 CC GENE (GlnA). P-II PREVENTS NR-II CATALYZED CONVERSION OF NR-I TO
 CC NR-I-PHOSPHATE, THE TRANSCRIPTION ACTIVATOR OF GlnA. WHEN P-II

CC IS URIDYLYLATED TO P-II-UMP. THESE EVENTS ARE REVERSED.
 CC -!- SUBUNIT: Homotrimer (By similarity).
 CC -!- SIMILARITY: Belongs to the P(II) protein family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X14012; CAA32177.1; -.
 CC PIR; S04377; S04377.
 CC HSP; P05826; 2PII.
 CC InterPro; IPR002187; PII_glnB.
 CC InterPro; IPR002332; PII_glnB_UMP_S.
 CC Pfam; PF00543; P-II; 1.
 CC PRINTS; PR00340; PIIGLNB.
 CC ProDom; PD001194; PII_glnB; 1.
 CC PROSITE; PS00496; PII_GlnB_UMP; 1.
 CC PROSITE; PS00638; PII_GlnB_CTER; 1.
 CC Transcription regulation; Nitrogen fixation.
 CC BINDING 51 51 UMP.
 CC MUTAGEN 50 50 E-Y: GLUTAMINE AUXOTROPHY.
 CC SEQUENCE 112 AA; 12429 MW; CCLSD58A2F225507 CRC64;
 CC -----
 CC Query Match 50.7%; Score 284.5; DB 1; Length 112;
 CC Best Local Similarity 50.9%; Pred. No. 5.6e-22;
 CC Matches 56; Conservative 27; Mismatches 26; Indels 1; Gaps 1;
 CC -----
 CC QY 3 KVEAIVRPRIQQVSSALLKIGIRGVTSVDRGFGAGGSTERHGGSEFSEDFKFAVKM 62
 CC Db 3 KIDALLKPKLDVREALAEGVGTGVTVEVKGFGKRGHTSYRGAEMVDFLPVKVI 61
 CC QY 63 EIVVKKQDQVESVINTIIEGATGIGDKIFLVPVSDVIRVTRGERGEKA 112
 CC Db 62 EIVVTDVDTVDVTIIRTAQKIGDKIFLVPVSDVIRVTRGERGEADA 111
 CC -----
 CC RESULT 10
 CC Y059 METJA STANDARD; PRT; 112 AA.
 CC AC Q60381;
 CC DT 01-NOV-1997 (Rel. 35, Created)
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE Hypothetical nitrogen regulatory PII-like protein MJ0059.
 CC GN M70059.
 CC OS Methanococcus jannaschii.
 CC OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 CC OC Methanocaldococcaceae; Methanocaldococcus.
 CC OX NCBI_TaxID=2190;
 CC RN [1]
 CC SEQUENCE FROM N.A.
 CC RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
 CC RX MEDLINE=96337999; PubMed=8689087;
 CC RA Bult C.J., White C., Olsen G.J., Zhou L., Fleischmann R.D.,
 CC Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
 CC Kervlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 CC Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 CC Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,
 CC Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 CC Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 CC Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 CC RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 CC jannaschii".
 CC RL Science 273:1058-1073 (1996).
 CC -!- SIMILARITY: Belongs to the P(II) protein family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U67464; AAB98041.1; -.
 CC PIR; C64307; C64307.
 CC HSP; P38504; LGNK.
 CC TIGR; MJ0059; -.
 CC InterPro; IPR002187; PII_glnB.
 CC InterPro; IPR002332; P-II; 1.
 CC Pfam; PF00543; P-II; 1.
 CC PRINTS; PR00340; PIIGLNB.
 CC ProDom; PD001194; PII_glnB; 1.
 CC PROSITE; PS00638; PII_GlnB_CTER; 1.
 CC PROSITE; PS00638; PII_GlnB_CTER; 1.
 CC Hypothetical protein; Transcription regulation; Complete proteome.
 CC BINDING 51 51 UMP (BY SIMILARITY).
 CC SEQUENCE 112 AA; 12490 MW; 3DC3F307C32FE9F3 CRC64;
 CC -----
 CC Query Match 50.4%; Score 282.5; DB 1; Length 112;
 CC Best Local Similarity 53.6%; Pred. No. 8.9e-22;
 CC Matches 59; Conservative 21; Mismatches 29; Indels 1; Gaps 1;
 CC -----
 CC QY 3 KVEAIVRPRIQQVSSALLKIGIRGVTSVDRGFGAGGSTERHGGSEFSEDFKFAVKM 62
 CC Db 3 KVEAIIIRPEKLEIVKKALS DAGYGVMTVSEVKGFGVGGIVYRGREIVD-LIPVKVI 61
 CC QY 63 EIVVKKQDQVESVINTIIEGATGIGDKIFLVPVSDVIRVTRGERGEKA 112
 CC Db 62 ELAVKESDVNDVIDIICENARTGNPGDKIFVIPVERVVRVTRKEGKEA 111
 CC -----
 CC RESULT 11
 CC GLNB PORPU STANDARD; PRT; 112 AA.
 CC ID GLNB PORPU
 CC AC P51254;
 CC DT 01-OCT-1996 (Rel. 34, Created)
 CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
 CC DT 01-OCT-1996 (Rel. 34, Last annotation update)
 CC DE Nitrogen regulatory protein P-II.
 CC GN GLNB.
 CC OS Porphyra purpurea.
 CC OC Chloroplast.
 CC OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
 CC OX NCBI_TaxID=2787;
 CC RN [1]
 CC SEQUENCE FROM N.A.
 CC RC STRAIN=Avonport;
 CC RA Reith M.E., Munnholland J.;
 CC RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast
 CC genome".
 CC RL Plant Mol. Biol. Rep. 13:333-335 (1995).
 CC CC -!- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE
 CC GLUTAMINE SYNTHETASE GENE (GLNA). P-II PREVENTS NR-II CATALYZED
 CC CONVERSION OF NR-I TO NR-I-PHOSPHATE, THE TRANSCRIPTIONAL
 CC ACTIVATOR OF GLNA. WHEN P-II IS URIDYLYLATED TO P-II-UMP, THESE
 CC EVENTS ARE REVERSED. WHEN THE RATIO OF GLN TO 2-KETOGLUTARATE
 CC DECREASES, P-II IS URIDYLYLATED TO P-II-UMP, WHICH CAUSES THE
 CC DEADENYLATION OF GLUTAMINE SYNTHETASE, SO ACTIVATING THE ENZYME
 CC (BY SIMILARITY).
 CC CC -!- SUBUNIT: Homotrimer (By similarity).
 CC CC -!- SIMILARITY: Belongs to the P(II) protein family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U38804; AAC08140.1; -.
 CC PIR; S73175; S73175.
 CC HSP; P05826; 2PII.

DR InterPro: IPR002187; PII_glnB.
DR InterPro: IPR002332; PII_glnB_UMP_S.
DR Pfam: PF00543; P-II; 1.
DR PRINTS: PR00340; PIIglnB.
DR ProDom: PD001194; PII_glnB; 1.
DR PROSITE: PS00496; PII_GlnB_UMP; 1.
DR PROSITE: PS00638; PII_GlnB_CTER; 1.
DR PROSITE: PS00638; PII_GlnB_CTER; 1.
KW Transcription regulation; Nitrogen fixation; Chloroplast.
FT BINDING 51 UMP (BY SIMILARITY).
SQ SEQUENCE 112 AA; 12320 MW; 29416AA49FC37A18 CRC64;

Query Match 49.8%; Score 279.5; DB 1; Length 112;
Best Local Similarity 51.8%; Pred. No. 1.8e-21;
Matches 57; Conservative 21; Mismatches 31; Indels 1; Gaps 1;

QY 3 KVEAIVRPRIQOVSSALLKIGIRGVTVSDVRGFGAGGSGTHERHGGSEFSEDKFVAKVM 62
DB 3 KIEAIRPFKLNEVLKLVKGGTGMVTVKVSFGGRKQGTERTYKSGEYSD-IIDKIKI 61
QY 63 EIVVKDQVESVINTIEGARTGEIGDKIFVLPVSDVIRVTRTGERGEKA 112
DB 62 EIIVSDKXKSITEIIKTAKTGEIGDKIFISDVEQVIRITNDLNSAA 111

RESULT 12
GLNB RHORU STANDARD; PRT; 112 AA.
ID GLNB RHORU STANDARD; PRT; 112 AA.
AC Q53044;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Nitrogen regulatory protein P-II.
GN GLNB.
OS Rhodospirillum rubrum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Rhodospirillaceae; Rhodospirillum.
OX NCBI_TaxID=1085;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 11170 / SL;
RX MEDLINE=96254013; PubMed=8704966;
RA Johansson M., Nordlund S.;
RT "Transcription of the glnB and glnA genes in the photosynthetic
bacterium Rhodospirillum rubrum.";
RL Microbiology 142:1265-1272(1996).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 11170 / SL;
RA Zhang Y., Ludden P.W., Roberts G.P.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE
GLUTAMINE SYNTHETASE GENE (GLNA). P-II PREVENTS NR-II CATALYZED
CONVERSION OF NR-I TO NR-I-PHOSPHATE, THE TRANSCRIPTIONAL
ACTIVATOR OF GLNA. WHEN P-II IS URIDYLATED TO P-II-UMP, THESE
EVENTS ARE REVERSED. WHEN THE RATIO OF GLN TO 2-KETOGLUTARATE
DECREASES, P-II IS URIDYLATED TO P-II-UMP, WHICH CAUSES THE
DEADENYLATION OF GLUTAMINE SYNTHETASE, SO ACTIVATING THE ENZYME.
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SIMILARITY: Belongs to the P(II) protein family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X84158; CAA58963.1; --
EMBL; AF029703; AAB84167.1; --
DR PIR; S52328; S52328.
DR HSSP; P05826; 2PII.
DR InterPro: IPR002187; PII_glnB.
DR PRINTS: PR00340; PIIglnB.
DR ProDom: PD001194; PII_glnB; 1.
DR PROSITE: PS00496; PII_GlnB_UMP; 1.
DR PROSITE: PS00638; PII_GlnB_CTER; 1.
KW Transcription regulation; Nitrogen fixation; Chloroplast.
FT BINDING 51 UMP (BY SIMILARITY).
SQ SEQUENCE 112 AA; 12320 MW; 29416AA49FC37A18 CRC64;

Query Match 49.8%; Score 279.5; DB 1; Length 112;
Best Local Similarity 51.8%; Pred. No. 1.8e-21;
Matches 57; Conservative 21; Mismatches 31; Indels 1; Gaps 1;

QY 3 KVEAIVRPRIQOVSSALLKIGIRGVTVSDVRGFGAGGSGTHERHGGSEFSEDKFVAKVM 62
DB 3 KIEAIRPFKLNEVLKLVKGGTGMVTVKVSFGGRKQGTERTYKSGEYSD-IIDKIKI 61
QY 63 EIVVKDQVESVINTIEGARTGEIGDKIFVLPVSDVIRVTRTGERGEKA 112
DB 62 EIIVSDKXKSITEIIKTAKTGEIGDKIFISDVEQVIRITNDLNSAA 111

RESULT 12
GLNB RHORU STANDARD; PRT; 112 AA.
ID GLNB RHORU STANDARD; PRT; 112 AA.
AC Q53044;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Nitrogen regulatory protein P-II.
GN GLNB.
OS Rhodospirillum rubrum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Rhodospirillaceae; Rhodospirillum.
OX NCBI_TaxID=1085;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 11170 / SL;
RX MEDLINE=96254013; PubMed=8704966;
RA Johansson M., Nordlund S.;
RT "Transcription of the glnB and glnA genes in the photosynthetic
bacterium Rhodospirillum rubrum.";
RL Microbiology 142:1265-1272(1996).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 11170 / SL;
RA Zhang Y., Ludden P.W., Roberts G.P.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE
GLUTAMINE SYNTHETASE GENE (GLNA). P-II PREVENTS NR-II CATALYZED
CONVERSION OF NR-I TO NR-I-PHOSPHATE, THE TRANSCRIPTIONAL
ACTIVATOR OF GLNA. WHEN P-II IS URIDYLATED TO P-II-UMP, THESE
EVENTS ARE REVERSED. WHEN THE RATIO OF GLN TO 2-KETOGLUTARATE
DECREASES, P-II IS URIDYLATED TO P-II-UMP, WHICH CAUSES THE
DEADENYLATION OF GLUTAMINE SYNTHETASE, SO ACTIVATING THE ENZYME.
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SIMILARITY: Belongs to the P(II) protein family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X84158; CAA58963.1; --
EMBL; AF029703; AAB84167.1; --
DR PIR; S52328; S52328.
DR HSSP; P05826; 2PII.
DR InterPro: IPR002187; PII_glnB.
DR PRINTS: PR00340; PIIglnB.
DR ProDom: PD001194; PII_glnB; 1.
DR PROSITE: PS00496; PII_GlnB_UMP; 1.
DR PROSITE: PS00638; PII_GlnB_CTER; 1.
KW Transcription regulation; Nitrogen fixation; Chloroplast.
FT BINDING 51 UMP (BY SIMILARITY).
SQ SEQUENCE 112 AA; 12320 MW; 29416AA49FC37A18 CRC64;

Query Match 49.8%; Score 279.5; DB 1; Length 112;
Best Local Similarity 51.8%; Pred. No. 1.8e-21;
Matches 57; Conservative 21; Mismatches 31; Indels 1; Gaps 1;

QY 3 KVEAIVRPRIQOVSSALLKIGIRGVTVSDVRGFGAGGSGTHERHGGSEFSEDKFVAKVM 62
DB 3 KIEAIRPFKLNEVLKLVKGGTGMVTVKVSFGGRKQGTERTYKSGEYSD-IIDKIKI 61
QY 63 EIVVKDQVESVINTIEGARTGEIGDKIFVLPVSDVIRVTRTGERGEKA 112
DB 62 EIIVSDKXKSITEIIKTAKTGEIGDKIFISDVEQVIRITNDLNSAA 111

RESULT 12
GLNB RHORU STANDARD; PRT; 112 AA.
ID GLNB RHORU STANDARD; PRT; 112 AA.
AC Q53044;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Nitrogen regulatory protein P-II.
GN GLNB.
OS Rhodospirillum rubrum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Rhodospirillaceae; Rhodospirillum.
OX NCBI_TaxID=1085;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 11170 / SL;
RX MEDLINE=96254013; PubMed=8704966;
RA Johansson M., Nordlund S.;
RT "Transcription of the glnB and glnA genes in the photosynthetic
bacterium Rhodospirillum rubrum.";
RL Microbiology 142:1265-1272(1996).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 11170 / SL;
RA Zhang Y., Ludden P.W., Roberts G.P.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE
GLUTAMINE SYNTHETASE GENE (GLNA). P-II PREVENTS NR-II CATALYZED
CONVERSION OF NR-I TO NR-I-PHOSPHATE, THE TRANSCRIPTIONAL
ACTIVATOR OF GLNA. WHEN P-II IS URIDYLATED TO P-II-UMP, THESE
EVENTS ARE REVERSED. WHEN THE RATIO OF GLN TO 2-KETOGLUTARATE
DECREASES, P-II IS URIDYLATED TO P-II-UMP, WHICH CAUSES THE
DEADENYLATION OF GLUTAMINE SYNTHETASE, SO ACTIVATING THE ENZYME.
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SIMILARITY: Belongs to the P(II) protein family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; X84158; CAA58963.1; --
EMBL; AF029703; AAB84167.1; --
DR PIR; S52328; S52328.
DR HSSP; P05826; 2PII.
DR InterPro: IPR002187; PII_glnB.
DR PRINTS: PR00340; PIIglnB.
DR ProDom: PD001194; PII_glnB; 1.
DR PROSITE: PS00496; PII_GlnB_UMP; 1.
DR PROSITE: PS00638; PII_GlnB_CTER; 1.
KW Transcription regulation; Nitrogen fixation; Chloroplast.
FT BINDING 51 UMP (BY SIMILARITY).
SQ SEQUENCE 112 AA; 12320 MW; 29416

Dd 3 KVEAIRPEKLEIVKKALSADAGVVGMTVSVKGRGVQGGIVERVRGREYVD-LIPVKVI 61

Qy 63 EIVVKKDQVBSEVINTEIGARTGEIDGKIIFVLFPVSDFIIRVTCTERG 109
 |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Dd 62 ELVVKEEDVNVIDLICENARTGNPGDKIIFVIPVERVVRKTKEEG 108
 |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:

RESULT 15

GLNB_RHIME STANDARD; PRT; 112 AA.

AC Q52905;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DE 28-FEB-2003 (Rel. 41, Last annotation update)

DE Nitrogen regulatory protein P-II.

GW GLNB OR R01639 OR SMC00947.

OS Rhizobium meliloti (Sinorhizobium meliloti).

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.

NB1_TaxID=382;
[1]

RN SEQUENCE FROM N.A.
RP STRAIN=RCR2011 / SU47;
RC MEDLINE=97085050; PubMed=8931324;
RX Arondeguy T., Huez I., Fournent J., Kahn D.;
RA "Symbiotic nitrogen fixation does not require adenylation of
RT glutamine synthetase I in Rhizobium meliloti.";
RL FEWS Microbiol. Lett. 145:33-40(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=1021;
RC MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubier F., Gouzy J.-J., Bothe G., Ampe F., Batut J.,
BO Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
PO Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebaud P., Vandebol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
Rt Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).

CC -!- FUNCTION: IN NITROGEN-LIMITING CONDITIONS, WHEN THE RATIO OF GLN
TO 2-KETOGLUTARATE DECREASES, P-II IS URIDYLATED TO P-II-UMP.
P-II-UMP ALLOWS THE DEADENYLATION OF GLUTAMINE SYNTHETASE (GS),
THIS ACTIVATING THE ENZYME. CONVERSELY, IN NITROGEN EXCESS P-II
IS DEURIDYLATED AND PROMOTES THE ADENYLATION OF GS.
-!- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE GS
GENE (GLNA). P-II PREVENTS NR-II CATALYZED CONVERSION OF NR-I TO
NR-I-PHOSPHATE, THE TRANSCRIPTIONAL ACTIVATOR OF GLNA. WHEN P-II
IS URIDYLATED TO P-II-UMP, THESE EVENTS ARE REVERSED.
CC -!- SUBUNIT: Homotrimer (by similarity).
CC -!- SIMILARITY: Belongs to the P(II) protein family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; U50385; AAC44623.1; --
DR EMBL; AL591787; CAC46218.1; --
DR HSSP; P05826; 2piI.
DR InterPro; IPR002187; PI1_glnB.
DR InterPro; IPR002332; PI1_GlnB_UMP_S.
DR Pfam; PF00543; P-II; 1.
DR PRINTS; PR00340; PIIGLB.
DR PRODOM; PD001194; PI1_glnB; 1.
DR PROSITE; PS00496; PI1_GLNB_UMP; 1.
DR PROSITE; PS00638; PI1_GLNB_UTER; 1.
KW Transcription regulation; Nitrogen fixation; Complete proteome.
FT BINDING 51 51 UMP (BY SIMILARITY).
SQ SEQUENCE 112 AA; 12281 MW; 85E6465EE6AB57001 CRC64;

Query Match 49.1%; Score 275.5; DB 1; Length 112;
Best Local Similarity 48.6%; Pred. No. 4.6e-21;
Matches 52; Conservative 27; Mismatches 27; Indels 1; Gaps 1;
QY 3 KVEAIVRPRIQOVSSALLKIGIRGVTVSDVRGGAQGGSTERRHGGSEFSEDKFVAKVKM 62
DB 3 KIEAIIKPKLDEVKEALQEVGLQGLTVTEAKGFGKQKGTLYRGAETVVD-FLEPKKV 61
QY 63 EIVVKKDQVESVINTIIEGARTGEIGDGKIFVLPSVDVIRVIRTERG 109
DB 62 EVVLADENAEAVTEALRNAAQTGRIGDGKIFVSNVEEVIRIRITGTG 108

Search completed: May 24, 2004, 19:13:34
Job time : 11.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 19:09:38 ; Search time 36 Seconds
(without alignments)
990.377 Million cell updates/sec

Title: US-09-756-541-1

Perfect score: 561

Sequence: 1 FYKVEAIVPRRIQQVSSAL.....VLPSDVIRVTRGERGEKAE 113

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriaph.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB ID	Description
1	561	130.0	196	10 Q9ZST4	Q9ZST4 arabidopsis
2	509	90.7	197	10 Q9ZST5	Q9ZST5 ricinus com
3	490	87.3	194	10 Q9ARI4	Q9ARI4 medicago sa
4	322.5	57.5	112	16 Q7V5R4	Q7V5R4 prochloroco
5	318.5	56.8	112	16 Q9L422	Q9L422 anabaena sp
6	317.5	56.6	112	16 Q8DLA5	Q8DLA5 synechococ
7	316.5	56.4	112	2 Q9LAQ3	Q9LAQ3 synechococ
8	308.5	55.0	112	16 Q9JWC4	Q9JWC4 neisseria m
9	307.5	54.8	112	16 Q9JXK6	Q9JXK6 neisseria m
10	305.5	54.5	112	2 Q8GQS4	Q8GQS4 acetobacter
11	304.5	54.3	112	16 Q7VEK6	Q7VEK6 haemophilus
12	303.5	54.1	112	2 Q9L400	Q9L400 prochloroco
13	303.5	54.1	112	16 Q7W025	Q7W025 prochloroco
14	301.5	53.7	133	16 Q7UWZ1	Q7UWZ1 rhodospirell
15	298.5	53.2	112	16 Q7VA51	Q7VA51 prochloroco
16	297.5	53.0	112	16 Q7U8Z7	Q7U8Z7 synechococ

17	296.5	52.9	112	2 Q9EZQ2	Q9ezq2 azoarcus sp
18	293.5	52.3	114	16 Q87MF6	Q87mf6 vibrio para
19	292.5	52.1	108	2 Q8KUJ2	Q8kuj2 neisseria m
20	292.5	52.1	114	16 Q8DBE2	Q8db2 vibrio vuln
21	291.5	52.0	112	16 Q8EBJ6	Q8ebj6 shewanella
22	289.5	51.6	114	16 Q9KFX3	Q9kpx3 vibrio chol
23	286.5	51.1	105	2 Q8KUJ5	Q8kuj5 neisseria m
24	285.5	50.9	112	2 Q31188	Q31188 rhodobacter
25	284.5	50.9	112	2 Q8ZCQ9	Q8zcg9 versinia pe
26	282.5	50.7	112	2 Q9AMW9	Q9amw9 azospirillu
27	281.5	50.2	112	16 Q8XWX5	Q8xwx5 ralstonia s
28	281.5	50.2	112	16 Q7WJ14	Q7wj14 bordetella
29	281.5	50.2	112	16 Q7WAE1	Q7wae1 bordetella
30	281.5	50.2	112	16 Q7VVG9	Q7vvg9 bordetella
31	278.5	49.6	112	16 Q9X705	Q9x705 corynebacte
32	276.5	49.3	112	2 P94125	P94125 azorhizobiu
33	276.5	49.3	112	16 Q8EAS3	Q8eas3 shewanella
34	276.5	49.3	112	17 Q56740	Q56740 methanococc
35	275.5	49.1	112	2 P94852	P94852 herbaspiril
36	275.5	49.1	157	16 Q8UEI8	Q8uei8 agrobacteri
37	274.5	48.9	112	16 Q8VH26	Q8vh26 brucella me
38	274.5	48.9	112	16 Q8G0T2	Q8got2 brucella su
39	273.5	48.8	101	2 Q8KHS5	Q8khs5 neisseria m
40	273.5	48.8	112	2 P70731	P70731 azospirillu
41	273.5	48.8	112	16 Q9A6W4	Q9a6w4 caulobacter
42	272.5	48.6	101	2 Q8KUJ0	Q8kuj0 neisseria m
43	272.5	48.6	112	16 Q8PP24	Q8fp24 corynebacte
44	272.5	48.6	121	16 Q9KUB6	Q9kub6 vibrio chol
45	271.5	48.4	100	2 Q8KI77	Q8ki77 neisseria m

ALIGNMENTS

RESULT 1

ID	Q9ZST4	PRELIMINARY;	PRT;	196 AA.
AC	Q9ZST4;			
DT	01-MAY-1999	(TRENBLrel. 10, Created)		
DT	01-MAY-1999	(TRENBLrel. 10, Last sequence update)		
DT	01-JUN-2003	(TRENBLrel. 24, Last annotation update)		
DE	PII protein (P II nitrogen sensing protein GUB I) (At4g01900).			
GN	T7B11.16 OR AT4G01900.			
OS	Arabidopsis thaliana (Mouse-ear cross).			
OC	Eukaryota; viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;			
OC	eurosids II; Brassicales; Brassicaceae; Arabidopsi.			
OX	NCBI_TaxID=3702;			
RN	[1]			
RN	SEQUENCE FROM N.A.			
RX	MEDLINE=99030678; PubMed=9811909;			
RA	Hsieh M.H., Lam H.M., van de Loo F.J., Coruzzi G.;			
RT	"A PII-like protein in Arabidopsis: putative role in nitrogen sensing.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 95:13965-13970 (1998).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Columbia;			
RA	Huang E.N., Nascimento L., de la Bastide M., Habermann K., Vil M.D.,			
RA	Preston R.R., Spiegel L.A., See L.H., Shah R., Matero A.,			
RA	O'Shaughnessy A., Rodriguez M., Shekher M., Swaby I., Schutz K.,			
RA	Parnell L.D., Dedhia N.N., McCombie W.R.;			
RT	"Arabidopsis thaliana BAC T7B11 from chromosome IV near 10 cm.";			
RL	Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,			
RA	Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,			
RA	Shekher M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,			
RA	Mewes H.W., Lemcke K., Mayer K.F.X.;			
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			

```

RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Shinn P., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Hsuan V.W.,
RA Ishida J., Jones T., Kaniya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis ORF clones.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF095455; AAC78333.1; -
DR EMBL; AC007138; AAD2652.1; -
DR DR EMBL; AL161493; CAB80683.1; -
DR DR EMBL; BT005209; AAO63273.1; -
DR DR PIR; D85024; D85024.
DR HSSP; P05826; 2PII.
DR GO; GO:030234; F:enzyme regulator activity; IEA.
DR GO; GO:0306806; P:regulation of nitrogen utilization; IEA.
DR InterPro; IPR002187; PII_glnb.
DR Pfam; PF00543; P-II; 1.
DR PRINTS; PRO0340; PIIGLNB.
DR ProDom; PD001194; PII_glnb; 1.
DR PROSITE; PS00638; PII_GLNb_CTER; 1.
DR SEQUENCE 196 AA; 21275 MW; FE740EA66776F157 CRC64;
SQ
Query Match 100.0%; Score 561; DB 10; Length 196;
Best Local Similarity 100.0%; Pred. No. 4.5e-48;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 FYKVEAIVRPWRIQQVSSALLKIGIRGVTVDVGRFGAQQGSTERHGGSEFSEDKFVAKV 60
DB 74 FYKVEAIVRPWRIQQVSSALLKIGIRGVTVDVGRFGAQQGSTERHGGSEFSEDKFVAKV 133
QY 61 KMEIVVKDQVESVINTIEGARTGIGDKIFLVPVSDVIRVTRGERGEKAE 113
DB 134 KMEIVVKDQVESVINTIEGARTGIGDKIFLVPVSDVIRVTRGERGEKAE 186
RESULT 2
ID Q9ZST5 PRELIMINARY; PRT; 197 AA.
AC Q9ZST5;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE PII protein (Fragment).
OS Ricinus communis (Castor bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids 2; Malpighiales; Euphorbiaceae; Acalyphoideae; Acalyphaceae;
OC Ricinus.
OX NCBI_TaxID=3988;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99030678; PubMed=9811909;
RA Hsieh M.H., Lam H.M., van de Loo F.J., Coruzzi G.;
RT "A PII-like protein in Arabidopsis: putative role in nitrogen sensing.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:13965-13970(1998).
DR HSSP; P05826; 2PII.
DR EMBL; AF095454; AAC78332.1; -
DR GO; GO:030234; F:enzyme regulator activity; IEA.
DR GO; GO:0306808; P:regulation of nitrogen utilization; IEA.
DR InterPro; IPR002187; PII_glnb.
DR Pfam; PF00543; P-II; 1.
DR PRINTS; PRO0340; PIIGLNB.
DR ProDom; PD001194; PII_glnb; 1.
DR PROSITE; PS00638; PII_GLNb_CTER; 1.
DR NON_TER 1
SQ SEQUENCE 197 AA; 21781 MW; 2D18A50F15406B20 CRC64;
Query Match 87.3%; Score 490; DB 10; Length 194;
Best Local Similarity 85.0%; Pred. No. 5.5e-41;
Matches 96; Conservative 10; Mismatches 7; Indels 0; Gaps 0;
QY 1 FYKVEAIVRPWRIQQVSSALLKIGIRGVTVDVGRFGAQQGSTERHGGSEFSEDKFVAKV 60
DB 70 FYKVEAIVRPWRIQQVSSALLKIGIRGVTVDVGRFGAQQGSTERHGGSEFSEDKFVAKV 129
QY 61 KMEIVVKDQVESVINTIEGARTGIGDKIFLVPVSDVIRVTRGERGEKAE 113
DB 130 KMEIVVKDQVEAVINKIMETARTGIGDKIFLVPVSDVIRVTRGERGEKAE 182
RESULT 4
ID Q7V5R4 PRELIMINARY; PRT; 112 AA.
AC Q7V5R4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nitrogen regulatory protein P-II.
GN GLNB OR PMT1481.
OS Prochlorococcus marinus (strain MIT 9313).
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=74547;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22825698; PubMed=12917642;
Query Match 90.7%; Score 503; DB 10; Length 197;
Best Local Similarity 90.3%; Pred. No. 7.1e-43;
Matches 102; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
QY 1 FYKVEAIVRPWRIQQVSSALLKIGIRGVTVDVGRFGAQQGSTERHGGSEFSEDKFVAKV 60
DB 72 FYKVEAIVRPWRIQQVSSALLKIGIRGVTVDVGRFGAQQGSTERHGGSEFSEDKFVAKV 131
QY 61 KMEIVVKDQVESVINTIEGARTGIGDKIFLVPVSDVIRVTRGERGEKAE 113
DB 132 KMEIVVKDQVEDVIEKITEARTGIGDKIFLVPVSDVIRVTRGERGEKAE 184
RESULT 3
ID Q9ARI4 PRELIMINARY; PRT; 194 AA.
AC Q9ARI4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE PII protein.
GN GLNB.
OS Medicago sativa (Alfalfa).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids 1; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
OX NCBI_TaxID=3879;
RN [1]
RP SEQUENCE FROM N.A.
RX Garcia-Ibáñeta D., Sengupta-Gopalan C.;
RT "Characterization of PII (GLNB) in alfalfa.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; A027892; AAK16221.1; -
DR HSSP; P38504; 1GNK.
DR GO; GO:030234; F:enzyme regulator activity; IEA.
DR GO; GO:0006808; P:regulation of nitrogen utilization; IEA.
DR InterPro; IPR002187; PII_glnb.
DR Pfam; PF00543; P-II; 1.
DR PRINTS; PRO0340; PIIGLNB.
DR ProDom; PD001194; PII_glnb; 1.
DR PROSITE; PS00638; PII_GLNb_CTER; 1.
DR SEQUENCE 194 AA; 21413 MW; 23FA623FF4D97450 CRC64;
Query Match 90.7%; Score 503; DB 10; Length 197;
Best Local Similarity 90.3%; Pred. No. 7.1e-43;
Matches 102; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
QY 1 FYKVEAIVRPWRIQQVSSALLKIGIRGVTVDVGRFGAQQGSTERHGGSEFSEDKFVAKV 60
DB 70 FYKVEAIVRPWRIQQVSSALLKIGIRGVTVDVGRFGAQQGSTERHGGSEFSEDKFVAKV 129
QY 61 KMEIVVKDQVESVINTIEGARTGIGDKIFLVPVSDVIRVTRGERGEKAE 113
DB 130 KMEIVVKDQVEAVINKIMETARTGIGDKIFLVPVSDVIRVTRGERGEKAE 182

```

RA	Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,
RA	Aligren N.A., Arellano A., Coleman M., Hauser I., Hess W.R.,
RA	Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,
RA	Shaw S.L., Seeglich C., Sullivan Y.B., Ting C.S., Tolonen A.,
RA	Webb E.A., Zinser E.R., Chisholm S.W.;
RT	"Genome divergence in two Prochlorococcus ecotypes reflects oceanic niche differentiation.";
RL	Nature 424:1042-1047(2003).
DR	EMBL; BX572099; CAE21656.1; --
KW	Complete proteome.
SQ	SEQUENCE 112 AA; 12362 MW; 5CA64D663ED3B65 CRC64;
Query Match	57.5%; Score 322.5; DB 16; Length 112;
Best Local Similarity	58.2%; Pred. No. 1.5e-24;
Matches	64; Conservative 21; Mismatches 24; Indels 1; Gaps 1;
Qy	3 KVEAIVRPWRIQVSSALLKIGIRGVTVSDVRGFGAGCGSTERRHGSGFSEDFKFAVKVM 62
	: : : : : : : : : : : : : : : : : : : :
Dd	3 KVEAIVRPFKLIEDVKLALVAELIIGTMTSVRGFRGRQGVRSSEFTVE-FLQLKV 61
	: : : : : : : : : : : : : : : : : : : :
Qy	63 EIUVKKQCVESVINTEIGARTGEIGDKIPVLPSVDVIRVRTGERKA 112
	: : : : : : : : : : : : : : : : : : : :
Dd	62 EVYYDDDKZAVNNAIAEAAKTGISGGIKFIISPDSVRIITGEROSKA 111
	: : : : : : : : : : : : : : : : : : : :

RESULT 5
Q9L422 PRELIMINARY; PRT; 112 AA.
ID Q9L422;
AC AC
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE PII protein (Nitrogen regulatory protein P-II).
GN GLNB OR ALL2319.
GN Anabaena sp. (strain PCC 7120).
OS Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OC NCBI_TaxID=133690;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PCC 7120;
RA Gonzalez L., Phalip V., Zhang C.C.;
RT "Phosphorylation of the signal transduction protein PII by the Ser/Thr
RT kinase PnKC in the cyanobacterium Anabaena sp. strain PCC 7120.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
DR EMBL; AJ251822; CAB75358.1; -
DR EMBL; AP003589; BAB74018.1; -
DR FIR; AH2095; AH2095.
DR HSSP; P05826; ZPII.
DR GO; GO:0030234; F:enzyme regulator activity; IEA.
DR GO; GO:0006908; F:regulation of nitrogen utilization; IEA.
DR InterPro; IPR002187; PII_GlnB.
DR InterPro; IPR002332; PII_GlnB_UMP_S.
DR Pfam; PF00543; P-II; 1.
DR PRINTS; PR00340; FIIGLN_B.
DR ProDom; PD001194; PII_GlnB; 1.
DR PROSITE; PS00638; PII_GlnB_CTER; 1.
DR PROSITE; PS00496; PII_GlnB_UMP; 1.
DR Complete proteome.
SQ SEQUENCE 112 AA; 12492 MW; F74E54C38B67596B CRC64;
Query Match 56.8%; Score 318.5; DB 16; Length 112;
Best Local Similarity 57.3%; Pred. No. 3.8e-24;

Fri May 28 14:57:01 2004

```

SEQUENCE FROM N.A.
RA STRAIN=PC7002;
RA Sakamoto T., Gruber T., Bryant D.A.;
RT "Nucleotide sequence of glnB and bcp genes in Synechococcus sp.
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF120107; AAF63031.1; -.
DR HSP; P05826; 2PIL.
DR GO; GO:0030234; F:enzyme regulator activity; IEA.
DR GO; GO:0006808; P:regulation of nitrogen utilization; IEA.
DR InterPro; IPR002187; PII_glnB.
DR InterPro; IPR002332; PII_glnB_UMP_S.
DR Pfam; PF00543; P-II; 1.
DR PRINTS; PR00340; PIIGLNB.
DR PRODOM; PD001194; PII_glnB; 1.
DR PROSITE; PS00638; PII_GlnB_CTER; 1.
DR PROSITE; PS00496; PII_GlnB_UMP; 1.
DR SEQUENCE 112 AA; 12432 MW; C3F0AA5625E9F870 CRC64;

Query Match 56.4%; Score 316.5; DB 2; Length 112;
Best Local Similarity 54.5%; Pred. No. 6.1e-24;
Matches 60; Conservative 27; Mismatches 22; Indels 1; Gaps 1;

QY 3 KVEAIVRPWRIQQVSSALLKIGIRGTVSDVVRGFGAQQGSTERHGGSEFSEDKFVAKVM 62
DB 3 KVEAIVRPFLDDEKIALVAGIVGTVSEVRGFGKQGTERTYRGSEYVE-FLQKLKI 61

63 EIVVKQDOVESVINTIIEGARTGEIGDKGKIFVLPSVDIVRVTGERGKA 112
62 EIVDDDDQVDAVDVKIVAARTGEIGDKGKIFSPVDQIVRITGEKDLA 111

SEQUENCE FROM N.A.
RA STRAIN=MC58 / Serogroup B;
RC MEDLINE=2017555; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Citterone H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Masiagnani V., Pizzo M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58.";
RL Science 287:1809-1815(2000).
DR EMBL; AE002548; AAF42322.1; -.
DR PIR; B81019; B81019.
DR HSP; P05826; 2PIL.
DR TIGR; NMB1995; -.
DR GO; GO:0030234; F:enzyme regulator activity; IEA.
DR GO; GO:0006808; P:regulation of nitrogen utilization; IEA.
DR InterPro; IPR002187; PII_glnB.
DR InterPro; IPR002332; PII_glnB_UMP_S.
DR Pfam; PF00543; P-II; 1.
DR PRINTS; PR00340; PIIGLNB.
DR PRODOM; PD001194; PII_glnB; 1.
DR PROSITE; PS00638; PII_GlnB_CTER; 1.
DR PROSITE; PS00496; PII_GlnB_UMP; 1.
DR SEQUENCE 112 AA; 12311 MW; 22963CF2526D4332 CRC64;

Query Match 54.8%; Score 307.5; DB 16; Length 112;
Best Local Similarity 54.5%; Pred. No. 4.8e-23;
Matches 60; Conservative 25; Mismatches 24; Indels 1; Gaps 1;

QY 3 KVEAIVRPWRIQQVSSALLKIGIRGTVSDVVRGFGAQQGSTERHGGSEFSEDKFVAKVM 62
DB 3 KIEAIVKPFKLDVREALTETIGTMTVSEVRGFGKQGTERTYRGSEYVE-FLPKIKI 61

63 EIVVKQDOVESVINTIIEGARTGEIGDKGKIFVLPSVDIVRVTGERGKA 112
62 EIVLADDAVERAIDVIVEVARSGKIGDKGKIFLVPVEAIRITGERDAA 111

SEQUENCE FROM N.A.
RA STRAIN=MC58 / Serogroup B;
RC MEDLINE=2017555; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Citterone H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Masiagnani V., Pizzo M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58.";
RL Science 287:1809-1815(2000).
DR EMBL; AE002548; AAF42322.1; -.
DR PIR; B81019; B81019.
DR HSP; P05826; 2PIL.
DR TIGR; NMB1995; -.
DR GO; GO:0030234; F:enzyme regulator activity; IEA.
DR GO; GO:0006808; P:regulation of nitrogen utilization; IEA.
DR InterPro; IPR002187; PII_glnB.
DR InterPro; IPR002332; PII_glnB_UMP_S.
DR Pfam; PF00543; P-II; 1.
DR PRINTS; PR00340; PIIGLNB.
DR PRODOM; PD001194; PII_glnB; 1.
DR PROSITE; PS00638; PII_GlnB_CTER; 1.
DR PROSITE; PS00496; PII_GlnB_UMP; 1.
DR SEQUENCE 112 AA; 12281 MW; CD893CF24A1A485F CRC64;

Query Match 56.4%; Score 316.5; DB 2; Length 112;
Best Local Similarity 54.5%; Pred. No. 6.1e-24;
Matches 60; Conservative 27; Mismatches 22; Indels 1; Gaps 1;

QY 3 KVEAIVRPWRIQQVSSALLKIGIRGTVSDVVRGFGAQQGSTERHGGSEFSEDKFVAKVM 62
DB 3 KVEAIVRPFLDDEKIALVAGIVGTVSEVRGFGKQGTERTYRGSEYVE-FLQKLKI 61

63 EIVVKQDOVESVINTIIEGARTGEIGDKGKIFVLPSVDIVRVTGERGKA 112
62 EIVDDDDQVDAVDVKIVAARTGEIGDKGKIFSPVDQIVRITGEKDLA 111

SEQUENCE FROM N.A.
RA STRAIN=22491 / Serogroup A / Serotype 4A;
RC MEDLINE=2022556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Baaham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis 22491.";
RL Nature 404:502-506(2000).
DR EMBL; AL162753; CAB83744.1; -.
DR PIR; H81961; H81961.
DR HSP; P05826; 2PIL.
DR GO; GO:0030234; F:enzyme regulator activity; IEA.
DR GO; GO:0006808; P:regulation of nitrogen utilization; IEA.
DR InterPro; IPR002187; PII_glnB.
DR InterPro; IPR002332; PII_glnB_UMP_S.
DR Pfam; PF00543; P-II; 1.
DR PRINTS; PR00340; PIIGLNB.
DR PRODOM; PD001194; PII_glnB; 1.
DR PROSITE; PS00638; PII_GlnB_CTER; 1.
DR PROSITE; PS00496; PII_GlnB_UMP; 1.
DR SEQUENCE 112 AA; 12281 MW; CD893CF24A1A485F CRC64;

```

Q8QC34;
01-MAR-2003 (Tremblrel. 23, Created)
01-MAR-2003 (Tremblrel. 23, Last sequence update)
01-OCT-2003 (Tremblrel. 25, Last annotation update)
Regulatory protein P-II.
GN GlnB.
OS Acetobacter diazotrophicus (Gluconacetobacter diazotrophicus).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Acetobacteraceae; Gluconacetobacter.
OX NCBI_TaxID=33396;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PAL5;
RA Perlova O., Meletzus D.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF318039; AAN59757.1; -;
DR GO: GO:000234; F:enzyme regulator activity; IEA.
DR GO: GO:0006808; P:regulation of nitrogen utilization; IEA.
DR InterPro: IPR002187; PII_GlnB.
DR InterPro: IPR002332; PII_GlnB_UMP_S.
DR Pfam: PF00543; P-II; 1.
DR PRINTS: PR00340; PII_GlnB.
DR PRODOM: PD001194; PII_GlnB; 1.
DR PROSITE: PS00638; PII_GlnB_CTER; 1.
DR PROSITE: PS00496; PII_GlnB_UMP; 1.
SQ SEQUENCE 112 AA; 12431 MW; 52AB6F9C0CC8D124 CRC64;
Query Match 54.5%; Score 305.5; DB 2; Length 112;
Best Local Similarity 52.7%; Pred. No. 7.6e-23;
Matches 58; Conservative 25; Mismatches 26; Indels 1; Gaps 1;
QY 3 KVEAIVRPWRIQQVSSALLKIGIRGVTVDVVRGFGAGGSTERHSGSEFSDKFKVAKVM 62
DB 3 KIEAIIKPKLDEKALHEIGLMTVTTEAKGFGKQKHTLYRGAETVD-FLPKVKL 61
QY 63 EIVVKQDQVESVINTIIEGARTGEIGDKIFVLPSDVIRVIRGGERKA 112
DB 62 EIVCADNLVDRAVETITMAAARTGRIGDKIFILPVEDVIRIRTGHEGSEA 111
RESULT 11
QY Q7VPK6 PRELIMINARY; PRT; 112 AA.
AC Q7VPK6;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Putative nitrogen regulatory protein P-II.
GN GlnB OR HD0062.
OS Haemophilus ducreyi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=730;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=35000HP / ATCC 700724;
RA Munson R.S. Jr., Ray W.C., Mahairas G., Sabo P., Mungur R.,
RA Johnson L., Nguyen D., Wang J., Forst C., Hood L.;
RA "The complete genome sequence of Haemophilus ducreyi."
RT Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL: AE017151; AAP95074.1; -;
RW Complete proteome.
SQ SEQUENCE 112 AA; 12570 MW; 9977658EA4864C90 CRC64;
Query Match 54.3%; Score 304.5; DB 16; Length 112;
Best Local Similarity 53.6%; Pred. No. 9.6e-23;
Matches 59; Conservative 25; Mismatches 25; Indels 1; Gaps 1;
QY 3 KVEAIVRPWRIQQVSSALLKIGIRGVTVDVVRGFGAGGSTERHSGSEFSDKFKVAKVM 62
DB 3 KIEAIIKPKLDDVREALTDVIGTMTVNEVKFGKQKHTLYRGAETVD-FLPKVKL 61
QY 63 EIVVKQDQVESVINTIIEGARTGEIGDKIFVLPSDVIRVIRGGERKA 112

Db 62 EIIIVDEQVESCNALIEATQTKIGDKIFVYDVVRVIRTGGEREA 111
RESULT 12
QY Q9L400 PRELIMINARY; PRT; 112 AA.
AC Q9L400;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE GlnB protein.
GN GlnB.
OS Prochlorococcus sp. (strain PCC 9511).
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=100363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=strain PCC 9511;
RA Fallinska K.A., Loiseau-de Goer S., Blondel A., Castets A.M.,
RA Rippka R., Tandeau de Marsac N.;
RT "In spite of synthesizing a cyanobacterial-type PII protein, the
RT oceanic photosynthetic prokaryote Prochlorococcus marinus, strain PCC
RT 9511, is unable to utilize nitrate."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ271089; CB87556.1; -;
DR HSP; P38504; 1GNK.
DR GO: GO:0030234; F:enzyme regulator activity; IEA.
DR GO: GO:0006808; P:regulation of nitrogen utilization; IEA.
DR InterPro: IPR002187; PII_GlnB.
DR Pfam: PF00543; P-II; 1.
DR PRINTS: PR00340; PII_GlnB.
DR PRODOM: PD001194; PII_GlnB; 1.
DR PROSITE: PS00638; PII_GlnB_CTER; 1.
SQ SEQUENCE 112 AA; 12314 MW; 42B3497E6B4472CD CRC64;
Query Match 54.1%; Score 303.5; DB 2; Length 112;
Best Local Similarity 52.7%; Pred. No. 1.2e-22;
Matches 58; Conservative 27; Mismatches 24; Indels 1; Gaps 1;
QY 3 KVEAIVRPWRIQQVSSALLKIGIRGVTVDVVRGFGAGGSTERHSGSEFSDKFKVAKVM 62
DB 3 KIEAIIKPKLEDDVKIALVNSGIVGTVSEVRGFGKQKHTLYRGAETVD-FLQKLKV 61
QY 63 EIVVKQDQVESVINTIIEGARTGEIGDKIFVLPSDVIRVIRGGERKA 112
DB 62 EVVVENKSVSDIAIAAATAKTEIGDKIFISSIDSVVIRITGDTDEEA 111
RESULT 13
QY Q7V025 PRELIMINARY; PRT; 112 AA.
AC Q7V025;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Nitrogen regulatory protein P-II.
GN GlnB OR PM1463.
OS Prochlorococcus marinus subsp. pastoris (strain CCMP 1378 / MED4).
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=59919;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=22825698; PubMed=12917642;
RA Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,
RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
RA Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,
RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
RA Webb E.A., Zenger E.R., Chisholm S.W.;
RT "Genome divergence in two Prochlorococcus ecotypes reflects oceanic
RT niche differentiation.";

RL Nature 424:1042-1047(2003).
DR EMBL; BX572094; CAE19922.1; -.
KW Complete proteome.
SQ SEQUENCE 112 AA; 12314 MW; 42B3497E6B4472CD CRC64;
Query Match 54.1%; Score 301.5; DB 16; Length 112;
Best Local Similarity 52.7%; Pred. No. 1.2e-22;
Matches 56; Conservative 27; Mismatches 24; Indels 1; Gaps 1;
QY 3 KVEAIVRPMRIQQVSSALLKIGIRGVTVDVVRGFGAGGSTERHGGSEFSEDKFVAKVKM 62
Db 3 KIEALIRPFKLEVDKIALVNLGIVGMTVSEVRGFGKQGVRYRGSEFTVE-FLQKLKV 61
QY 63 EIVVKKDQVESVINTIEGARTGEIGDKIFVLPVSDVIRVIRGTGERGEKA 112
Db 62 EVVVENKSVVIDATAEAARKTGEIGDKIFISSIDSVVIRITGDTDEEA 111
RESULT 14
Q7U0Z1 PRELIMINARY; PRT; 133 AA.
ID Q7U0Z1
AC Q7U0Z1;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nitrogen regulatory protein P-II.
GN Rhodopirellula baltica.
OS Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
Schlesner H., Amann R., Reinhardt R.;
RT *Complete genome sequence of the marine planctomycete Pirellula sp.
strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294138; CAD72936.1; -.
KW Complete proteome.
SQ SEQUENCE 133 AA; 14941 MW; 05D489DB4D7B169D CRC64;
Query Match 53.7%; Score 301.5; DB 16; Length 133;
Best Local Similarity 54.5%; Pred. No. 2.4e-22;
Matches 60; Conservative 27; Mismatches 22; Indels 1; Gaps 1;
QY 3 KVEAIVRPMRIQQVSSALLKIGIRGVTVDVVRGFGAGGSTERHGGSEFSEDKFVAKVKM 62
Db 24 KVEAIVRHFHKLIEDVKNALTEQGHGTMVSEVRGFGKQGVRYRGSEFTVE-FVPRVKI 82
QY 63 EIVVKKDQVESVINTIEGARTGEIGDKIFVLPVSDVIRVIRGTGERGEKA 112
Db 83 EVVCTSDNLQIVDITLQTAQTGIGDKIFVFNLEESVIRITGERGEKA 132
RESULT 15
Q7VA51 PRELIMINARY; PRT; 112 AA.
ID Q7VA51
AC Q7VA51;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nitrogen regulatory protein PII.
GN GLNK OR Prol1616.
OS Prochlorococcus marinus.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
CX NCBI_TaxID=1219;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=GARG / COMP 1375 / SSI120;
RX MEDLINE=22810154; PubMed=12917486;
RA Dufresne A., Salanoubat M., Partensky F., Artiguenave F., Axmann I.M.,
Barbe V., Duprat S., Galperin M.Y., Koonin E.V., Le Gall F.,
Makarova K.S., Ostrowski M., Oztas S., Robert C., Rogozin I.B.,
Scanlan D.J., Tandeau de Marsac N., Weissenbach J., Wincker P.,
Wolf Y.I., Hess W.R.;
RT *Genome sequence of the cyanobacterium Prochlorococcus marinus SSI20,
a nearly minimal oxyphototrophic genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:10020-10025(2003).
DR EMBL; AB017165; AAQ00660.1; -.
KW Complete proteome.
SQ SEQUENCE 112 AA; 12239 MW; 3C4B9F51F49806EF CRC64;
Query Match 53.2%; Score 298.5; DB 16; Length 112;
Best Local Similarity 52.7%; Pred. No. 3.8e-22;
Matches 56; Conservative 24; Mismatches 27; Indels 1; Gaps 1;
QY 3 KVEAIVRPMRIQQVSSALLKIGIRGVTVDVVRGFGAGGSTERHGGSEFSEDKFVAKVKM 62
Db 3 KIEALIRPFKLEVDKIALVNLGIVGMTVSEVRGFGKQGVRYRGSEFTVE-FLQKLKI 61
QY 63 EIVVKKDQVESVINTIEGARTGEIGDKIFVLPVSDVIRVIRGTGERGEKA 112
Db 62 EVVVADESVDGVINKAIAEAARKTGEIGDKIFVSSIEFVLRITGESDDSA 111
Search completed: May 24, 2004, 19:15:01
Job time : 38 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 19:05:42 ; Search time 51 Seconds
(without alignments)
626.037 Million cell updates/sec

Title: US-09-756-541-2
Perfect score: 557
Sequence: 1 FYKVEAILRPWRVQVSSAL.....LLPVSDEVIRVTGERGDAE 113

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: geneseqp1980s:*

2: geneseqp1990s:*

3: geneseqp2000s:*

4: geneseqp2001s:*

5: geneseqp2002s:*

6: geneseqp2003as:*

7: geneseqp2003bs:*

8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	557	100.0	113	4	AAB69496 Ricinus c
2	509	91.4	113	4	AAB69495 Arabidops
3	306.5	55.0	112	4	AAB69503 Synethoco
4	303.5	54.5	121	6	ABP79052 N. gonorr
5	281.5	50.5	112	4	AAB69501 Azospiril
6	276.5	49.6	112	4	AAB69497 Klebsiell
7	274.5	49.3	112	4	AAB69498 Escherich
8	270.5	48.6	112	4	Agg93226 C glutami
9	256.5	46.1	124	6	ADA34150 Acinetoba
10	255.5	45.9	112	4	AAB69502 Rhodobact
11	251	45.1	111	4	AAB69499 Rhizobium
12	248	44.5	111	4	AAB69500 Bradyrhiz
13	245.5	44.1	160	4	ABG24254 Novel hum
14	240.5	43.2	89	4	ABG79181 Corynebac
15	218	39.1	112	5	ABBS4937 Lactococc
16	215.5	38.7	91	5	ABP03233 Human ORP
17	207.5	37.3	121	5	ABB49279 Listeria
18	183	32.9	243	4	ABG24253 Novel hum
19	183	32.9	579	4	ABG24250 Novel hum
20	178	32.0	111	4	ABG69505 Methanoco
21	159	28.5	1285	4	ABG26219 Novel hum
22	153.5	27.6	105	4	AAB69504 Methanoco
23	120	21.5	736	4	ABG24252 Novel hum
24	84.5	15.2	1036	6	ABU49254 Protein e
25	78	14.0	639	4	ABG21490 Novel hum

26	77	13.8	709	5	ABP30163 Streptoco
27	77	13.8	713	5	ABP25678 Streptoco
28	75.5	13.6	823	6	ABU21334 Protein e
29	74.5	13.4	274	4	ABB69532 Drosophil
30	73.5	13.2	374	2	AA43982 Horse alc
31	73.5	13.2	374	2	AA43981 Horse alc
32	73.5	13.2	547	2	AAR67385 Mitochond
33	73.5	13.2	554	5	ABP28230 Streptoco
34	72	12.9	169	5	ABP41194 Human ova
35	72	12.9	187	7	ABE24811 Human kin
36	71.5	12.8	541	6	ABU49927 Protein e
37	71	12.7	773	5	ABBS5204 Lactococc
38	70.5	12.7	573	2	AAR04713 Amino aci
39	70.5	12.7	573	2	AAW01657 Human hea
40	70.5	12.7	573	2	AAW14946 Human hea
41	70.5	12.7	573	2	AAW12345 Human hea
42	70.5	12.7	573	2	AAW56120 Protein s
43	70.5	12.7	573	2	AAW23926 Amino aci
44	70.5	12.7	573	3	AAW93334 Amino aci
45	70.5	12.7	573	4	AAE11756 Human hea

ALIGNMENTS

RESULT 1
AAB69496
ID AAB69496 standard; protein; 113 AA.
XX
AC AAB69496;
XX
23-APR-2001 (first entry)
XX
Ricinun communis PII protein.
XX
Castor bean; PII; plant nitrogen regulatory gene; P-PII;
KW nitrogen assimilation; transgenic plant; herbicide screening.
XX
Ricinun communis.
XX
US6177275-B1.
XX
23-JAN-2001.
XX
23-JUL-1997; 97US-00899330.
XX
24-JUL-1996; 96US-0022328P.
XX
(UUNY) UNIV NEW YORK STATE.
XX
Coruzzi GM, Lam H, Hsieh M;
XX
WPI; 2001-158572/16.
XX
N-PSDB; AAF58582, AAF58584.
XX
Novel P-PII genes capable of regulating plant nitrogen assimilation,
PT useful for transgenic plant production, and as probes for isolating
XX additional genomic clones having P-PII gene promoters.
XX
Claim 1; Fig 1; 35pp; English.
XX
The present sequence is encoded by a nitrogen regulatory PII gene. Novel
XX plant PII (also called P-PII) nucleotide sequences have been isolated.
XX They are useful for regulating nitrogen assimilation in plants, and in
XX transgenic plant production. They are also used to engineer organisms
XX that overexpress wild-type or mutant P-PII regulatory proteins. P-PII
XX proteins are useful for in vitro screening of herbicides. P-PII
XX nucleotides may be used as probes for isolating additional genomic clones
XX with the promoters of P-PII genes. P-PII promoters are light- and/or
XX sucrose-inducible, and are suitable for genetic engineering of plants
XX
XX Sequence 113 AA;

Query Match 100.0%; Score 557; DB 4; Length 113;
 Best Local Similarity 100.0%; Pred. No. 1.2e-57;
 Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FYKVEAILRPWRVSVSSALLKIGIRGVTVSDVVRGFGAQQGSTERQGGSEDFKFAKV 60
 DB 1 FYKVEAILRPWRVSVSSALLKIGIRGVTVSDVVRGFGAQQGSTERQGGSEDFKFAKV 60

QY 61 KMEIIVSKDQVEDVIEKIEEARTGEIGDGKIFLLPVSDVIRVTRGGRDCAE 113
 DB 61 KMEIIVSKDQVEDVIEKIEEARTGEIGDGKIFLLPVSDVIRVTRGGRDCAE 113

RESULT 2
 AAB69495
 ID AAB69495 standard; protein; 113 AA.
 XX
 AC AAB69495;
 XX
 DT 23-APR-2001 (first entry)
 XX
 DE Arabidopsis thaliana PII protein.
 XX
 KW Arabidopsis thaliana; PII; plant nitrogen regulatory gene; P-II;
 KW nitrogen assimilation; transgenic plant; herbicide screening.
 XX
 OS Arabidopsis thaliana.
 XX
 PN US6177275-B1.
 XX
 PD 23-JAN-2001.
 XX
 PF 23-JUL-1997; 97US-00899330.
 XX
 PR 24-JUL-1996; 96US-0022328P.
 XX
 PA (UNY) UNIV NEW YORK STATE.
 XX
 PI Coruzzi GM, Lam H, Heieh M;
 XX
 DR WPI; 2001-158572/16.
 XX
 PT Novel P-II genes capable of regulating plant nitrogen assimilation,
 PT useful for transgenic plant production, and as probes for isolating
 PT additional genomic clones having P-II gene promoters.
 XX
 PS Claim 1; Fig 1; 35pp; English.
 XX
 CC The present sequence is encoded by a nitrogen regulatory PII gene. Novel
 CC plant PII (also called P-II) nucleotide sequences have been isolated.
 CC They are useful for regulating nitrogen assimilation in plants, and in
 CC transgenic plant production. They are also used to engineer organisms
 CC that overexpress wild-type or mutant P-II regulatory proteins. P-II
 CC proteins are useful for in vitro screening of herbicides. P-II
 CC nucleotides may be used as probes for isolating additional genomic clones
 CC with the promoters of P-II genes. P-II promoters are light- and/or
 CC sucrose-inducible, and are suitable for genetic engineering of plants

Query Match 55.0%; Score 306.5; DB 4; Length 112;
 Best Local Similarity 56.4%; Pred. No. 4.2e-28;
 Matches 62; Conservative 21; Mismatches 26; Indels 1; Gaps 1;

QY 3 KVEAILRPWRVSVSSALLKIGIRGVTVSDVVRGFGAQQGSTERQGGSEDFKFAKV 62
 DB 3 KVEAILRPWRVSVSSALLKIGIRGVTVSDVVRGFGAQQGSTERQGGSEDFKFAKV 62

QY 63 EIVVSKDQVEDVIEKIEEARTGEIGDGKIFLLPVSDVIRVTRGGRDCAE 112
 DB 63 EIVVSKDQVEDVIEKIEEARTGEIGDGKIFLLPVSDVIRVTRGGRDCAE 112

RESULT 4
 AAB79052
 ID AAB79052 standard; protein; 121 AA.
 XX
 AC AAB79052;
 XX
 DT 07-MAR-2003 (first entry)
 XX
 DE N. gonorrhoeae amino acid sequence SEQ ID 4634.
 XX
 KW Antibacterial; infection; vaccine; gene therapy.
 XX

OS Neisseria gonorrhoeae.
 XX WO200279243-A2.
 XX 10-OCT-2002.
 XX 12-FEB-2002; 2002WO-IB002069.
 XX 12-FEB-2001; 2001GB-00003424.
 XX (CHIR-) CHIRON SPA.
 XX Fontana MR, Pizza M, Masignani V, Monaci E;
 XX WPI; 2003-058415/05.
 XX N-PSDB; AB240022.
 XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a
 XX medicament for treating or preventing N. gonorrhoeae infection.
 XX Disclosure; Page 528; 815pp; English.
 XX The present invention relates to proteins from Neisseria gonorrhoeae.
 XX Also disclosed are the nucleic acid molecules encoding the proteins and
 XX antibodies that specifically bind to the proteins. The composition
 XX comprising the protein, nucleic acid or antibody is useful for the
 XX manufacture of a medicament for treating or preventing N. gonorrhoeae
 XX infection, this may be in the form of a vaccine or gene therapy.
 XX Sequences given in records ABP76736-ABP81046 represent nucleic acid
 XX molecules of the invention
 XX Sequence 121 AA;
 SQ Query Match 54.5%; Score 303.5; DB 6; Length 121;
 Best Local Similarity 54.5%; Pred. No. 1.1e-27;
 Matches 60; Conservative 25; Mismatches 24; Indels 1; Gaps 1;
 QY 3 KVEAILRPNRVSVSSALLKIGIRGVTVDVDFGAGGSTERQGGSEFSEDFKFAKVM 62
 Db 12 KTEAIVKPKLDDREALTEIGITGMTVSEVKGFGKQGHTEYRGAEXAVD-FLPKVKI 70
 QY 63 EIVVSKQVEDVIEKIEEARTGEIGDKIFLLPVSDVIRVTRGERGDKA 112
 Db 71 ELVLADDAVERAIDVIVEVARSGKIGDKIFVLPVPEAIRTRGERSDAA 120
 RESULT 5
 AAB69501
 ID AAB69501 standard; protein; 112 AA.
 AC AAB69501;
 DT 11-SEP-2003 (revised)
 DT 23-APR-2001 (first entry)
 XX Azospirillum brasiliense PII protein.
 XX Azospirillum brasiliense; PII; plant nitrogen regulatory gene; P-II;
 XX nitrogen assimilation; transgenic plant; herbicide screening.
 XX Azospirillum brasiliense.
 XX US6177275-B1.
 XX 23-JAN-2001.
 XX 23-JUL-1997; 97US-00899330.
 XX 24-JUL-1996; 96US-0022328P.
 XX (UUNY) UNIV NEW YORK STATE.
 XX Coruzzi GM, Lam H, Hsieh M;
 XX WPI; 2001-158572/16.
 XX Novel P-II genes capable of regulating plant nitrogen assimilation,
 XX useful for transgenic plant production, and as probes for isolating
 XX additional genomic clones having P-II gene promoters.
 XX Example; Fig 1; 35pp; English.
 XX The present sequence is encoded by a nitrogen regulatory PII gene. Novel
 XX plant PII (also called P-II) nucleotide sequences have been isolated.
 XX They are useful for regulating nitrogen assimilation in plants, and in
 XX transgenic plant production. They are also used to engineer organisms
 XX that overexpress wild-type or mutant P-II regulatory proteins. P-II
 XX proteins are useful for in vitro screening of herbicides. P-II
 XX nucleotides may be used as probes for isolating additional genomic clones
 XX with the promoters of P-II genes. P-II promoters are light- and/or
 XX sucrose-inducible, and are suitable for genetic engineering of plants.
 XX (Updated on 11-SEP-2003 to standardise OS field)
 XX Sequence 112 AA;
 SQ Query Match 50.5%; Score 281.5; DB 4; Length 112;
 Best Local Similarity 49.1%; Pred. No. 3.8e-25;
 Matches 54; Conservative 26; Mismatches 29; Indels 1; Gaps 1;
 QY 3 KVEAILRPNRVSVSSALLKIGIRGVTVDVDFGAGGSTERQGGSEFSEDFKFAKVM 62
 Db 3 KIEAIIKPKFLDVEKALHEVGIKGTITVTEAKGFGKQGHTEYRGAEXAVD-FLPKVKI 61
 QY 63 EIVVSKQVEDVIEKIEEARTGEIGDKIFLLPVSDVIRVTRGERGDKA 112
 Db 62 EVMEDSLVERAIEAQQAAHTGRIGDKIFVTPVEEVVIRTRGEGGDA 111
 RESULT 6
 AAB69497
 ID AAB69497 standard; protein; 112 AA.
 AC AAB69497;
 DT 23-APR-2001 (first entry)
 XX Klebsiella pneumoniae PII protein.
 XX Klebsiella pneumoniae; PII; plant nitrogen regulatory gene; P-II;
 XX nitrogen assimilation; transgenic plant; herbicide screening.
 XX Klebsiella pneumoniae.
 XX US6177275-B1.
 XX 23-JAN-2001.
 XX 23-JUL-1997; 97US-00899330.
 XX 24-JUL-1996; 96US-0022328P.
 XX (UUNY) UNIV NEW YORK STATE.
 XX Coruzzi GM, Lam H, Hsieh M;
 XX WPI; 2001-158572/16.
 XX Novel P-II genes capable of regulating plant nitrogen assimilation,
 XX useful for transgenic plant production, and as probes for isolating
 XX additional genomic clones having P-II gene promoters.
 XX Example; Fig 1; 35pp; English.
 XX The present sequence is encoded by a nitrogen regulatory PII gene. Novel
 XX plant PII (also called P-II) nucleotide sequences have been isolated.
 XX They are useful for regulating nitrogen assimilation in plants, and in

XX WPI; 2001-158572/16.
 XX Novel P-II genes capable of regulating plant nitrogen assimilation,
 XX useful for transgenic plant production, and as probes for isolating
 XX additional genomic clones having P-II gene promoters.
 XX Example; Fig 1; 35pp; English.
 XX The present sequence is encoded by a nitrogen regulatory PII gene. Novel
 XX plant PII (also called P-II) nucleotide sequences have been isolated.
 XX They are useful for regulating nitrogen assimilation in plants, and in
 XX transgenic plant production. They are also used to engineer organisms
 XX that overexpress wild-type or mutant P-II regulatory proteins. P-II
 XX proteins are useful for in vitro screening of herbicides. P-II
 XX nucleotides may be used as probes for isolating additional genomic clones
 XX with the promoters of P-II genes. P-II promoters are light- and/or
 XX sucrose-inducible, and are suitable for genetic engineering of plants.
 XX (Updated on 11-SEP-2003 to standardise OS field)
 XX Sequence 112 AA;
 SQ Query Match 50.5%; Score 281.5; DB 4; Length 112;
 Best Local Similarity 49.1%; Pred. No. 3.8e-25;
 Matches 54; Conservative 26; Mismatches 29; Indels 1; Gaps 1;
 QY 3 KVEAILRPNRVSVSSALLKIGIRGVTVDVDFGAGGSTERQGGSEFSEDFKFAKVM 62
 Db 3 KIEAIIKPKFLDVEKALHEVGIKGTITVTEAKGFGKQGHTEYRGAEXAVD-FLPKVKI 61
 QY 63 EIVVSKQVEDVIEKIEEARTGEIGDKIFLLPVSDVIRVTRGERGDKA 112
 Db 62 EVMEDSLVERAIEAQQAAHTGRIGDKIFVTPVEEVVIRTRGEGGDA 111
 RESULT 6
 AAB69497
 ID AAB69497 standard; protein; 112 AA.
 AC AAB69497;
 DT 23-APR-2001 (first entry)
 XX Klebsiella pneumoniae PII protein.
 XX Klebsiella pneumoniae; PII; plant nitrogen regulatory gene; P-II;
 XX nitrogen assimilation; transgenic plant; herbicide screening.
 XX Klebsiella pneumoniae.
 XX US6177275-B1.
 XX 23-JAN-2001.
 XX 23-JUL-1997; 97US-00899330.
 XX 24-JUL-1996; 96US-0022328P.
 XX (UUNY) UNIV NEW YORK STATE.
 XX Coruzzi GM, Lam H, Hsieh M;
 XX WPI; 2001-158572/16.
 XX Novel P-II genes capable of regulating plant nitrogen assimilation,
 XX useful for transgenic plant production, and as probes for isolating
 XX additional genomic clones having P-II gene promoters.
 XX Example; Fig 1; 35pp; English.
 XX The present sequence is encoded by a nitrogen regulatory PII gene. Novel
 XX plant PII (also called P-II) nucleotide sequences have been isolated.
 XX They are useful for regulating nitrogen assimilation in plants, and in

transgenic plant production. They are also used to engineer organisms that overexpress wild-type or mutant P-Pil regulatory proteins. P-Pil proteins are useful for in vitro screening of herbicides. P-Pil nucleotides may be used as probes for isolating additional genomic clones with the promoters of P-pil genes. P-Pil promoters are light- and/or sucrose-inducible, and are suitable for genetic engineering of plants

Query Match	49.6%;	Score 276.5;	DB 4;	Length 112;
Best Local Similarity	50.0%;	Pred. No. 1.5e-24;		
Matches	55: Conservative	26: Mismatches	28: Indels	1: Gaps

Qy	3	KVERAILPRNRSQVSALLKIGIRVTGTVSDYRPFGAQGGSTERQGGSESESDKTFVAKVM	62
Qb	3	KIDAIIPKPKLDDVREALAEVGI TGMVTEVKGFGRQKGTHTLYRGAZYWD -FLP KVI	61

QY	63	EIVSKQVEDVIEKIIIEARTGEIGDKIFLLPVSDVIRVTRTGERGDKA	112
		: : :: : : : : :	
DB	62	EIVTDRIVDTCVYTIISTACTGKIGDKIFVFDVARVIRTRTGEEDDAA	111

RESULT 7
AAB69498
ID AAB69498 standard: protein; 112 AA.

AA	AAB69498;	
AC		
XX		
DT	23-APR-2001	(first entry)
XX		
DE	Escherichia coli PII protein.	

AA Escherichia coli; PII; plant nitrogen regulatory gene; P-PII;
KW nitrogen assimilation; transgenic plant; herbicide screening;
KW nitrogen assimilation; transgenic plant; herbicide screening;

XX Escherichia coli.
CS
XX
PN US6177275-B1.

XX	
PD	23-JAN-2001.
XX	
PF	23-JUL-1997: 97US-00899330.

XX 24-JUL-1996; 96US-0022328P.
PR
XX
PA (UYN) UNIV NEW YORK STATE.

XX Coruzzi GM, Lam H, Hsieh M;
XX
DR WPI: 2001-158572/16.

Novel P-II genes capable of regulating plant nitrogen assimilation, useful for transgenic plant production, and as probes for isolating additional genomic clones having P-II gene promoters.

PS Example; Fig 1; 35pp; English.

xx The present sequence is encoded by a nitrogen regulatory PII gene. Novel
cc plant PII (also called p-PII) nucleotide sequences have been isolated.
cc They are useful for regulating nitrogen assimilation in plants, and in
cc transgenic plant production. They are also used to engineer organisms
cc that overexpress wild-type or mutant p-PII regulatory proteins. p-PII
cc proteins are useful for *in vitro* screening of herbicides. p-PII
cc nucleotides may be used as probes for isolating additional genomic clones
cc with the promoters of p-PII genes. p-PII promoters are light- and/or
cc sucrose-inducible, and are suitable for genetic engineering of plants

Sequence 112 AA;

Query Match	49.3%	Score 274.5;	DB 4;	Length 112;
Best Local Similarity	50.0%;	Pred. No. 2.5e-24;		
Matches	55;	Conservative	25;	Mismatches 29;
				Indels 1

[illegible]

RESULT 8	
AAG93226	
ID	AAG93226 standard; protein; 112 AA.
XX	
XX	AAG93226;
XX	AC
XX	XX
XX	26-SEP-2001 (first entry)
XX	
DE	C glutamicum protein fragment SEQ ID NO: 6980.

XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis.
KW organic acid synthesis.

OS *Corynebacterium glutamicum*.

PN EP1108790-A2.

PD 20-JUN-2001.

18-DEC-2000; 2000EP-00127688.

PR 16-DEC-1999; 99JP-00377484-

PR 03-AUG-2000: 2000JP-00280988.

PA (RYOW) KYOWA HAKKO KOGYO KK.

AA Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;

WPI: 2001-376931/40.

DR N-PSDB: AAH68445.

Novel polynucleotides derived from Coryneform bacteria, for identifying mutation point of a gene, measuring expression of a gene, analyzing expression profile or pattern of a gene and identifying homologous gene.

xx
ps
Claim 29: SEO ID NO 6980: 246pp + Sequence Listing; English.

The present invention provides a number of nucleotide and protein sequences from the Coryneform bacterium *Corynebacterium glutamicum*. These are useful for identifying the mutation point of a gene derived from a mutant of coryneform bacterium, measuring expression amount and analysing the expression profile or expression pattern of a gene derived from *Coryneform bacterium*, and identifying a homologue of a gene derived from *Coryneform bacterium*. *Coryneform bacteria* are useful for producing amino acids, nucleic acids, vitamins, saccharides and organic acids, particularly L-lysine. The present sequence is a protein described in the exemplification of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the European Patent Office.

Sequence 112 AA:

Query Match: 48.6%; Score 270.5; DB 4; Length 112;

Best Local Similarity 45.0%; Rec: 30.7; 35.0
Matches 49; Conservative 33; Mismatches 26; Indels 1; Gaps 1

QY 4 VEAIIIPWRVYQVSSALLKIGIRGVTYSDDVRFGAQQGSTERQGSSEFSDKFKVAKVME 63

63 VIISDAQAEVINIIVETARTCKVGDGKVMNIEELVVRTGERGEAA 111

DE Rhodobacter capsulatus PII protein.
 XX Rhodobacter capsulatus; PII; plant nitrogen regulatory gene; P-II;
 KW nitrogen assimilation; transgenic plant; herbicide screening.
 XX Rhodobacter capsulatus.
 OS Rhodobacter capsulatus.
 XX US6177275-B1.
 PN 23-JAN-2001.
 XX 23-JUL-1997; 97US-00899330.
 XX 24-JUL-1996; 96US-0022328P.
 XX (UYN) UNIV NEW YORK STATE.
 XX Coruzzi GM, Lam H, Hsieh M;
 XX WPI; 2001-158572/16.
 XX Novel P-II genes capable of regulating plant nitrogen assimilation,
 PT useful for transgenic plant production, and as probes for isolating
 PT additional genomic clones having P-II gene promoters.
 XX Example; Fig 1; 35pp; English.
 XX The present sequence is encoded by a nitrogen regulatory PII gene. Novel
 CC plant PII (also called P-II) nucleotide sequences have been isolated.
 CC They are useful for regulating nitrogen assimilation in plants, and in
 CC transgenic plant production. They are also used to engineer organisms
 CC that overexpress wild-type or mutant P-II regulatory proteins. P-II
 CC proteins are useful for in vitro screening of herbicides. P-II
 CC nucleotides may be used as probes for isolating additional genomic clones
 CC with the promoters of P-II genes. P-II promoters are light- and/or
 CC sucrose-inducible, and are suitable for genetic engineering of plants
 XX SQ Sequence 112 AA;
 Query Match 45.9%; Score 255.5; DB 4; Length 112;
 Best Local Similarity 47.3%; Pred. No. 4.4e-22;
 Matches 52; Conservative 26; Mismatches 31; Indels 1; Gaps 1;
 QY 3 KVEAILRPRVSVSALLKIGIRGVTVSDVRGFGAGGSTERQGGSESEDFKVAKVM 62
 DB 3 KVEAILRPRVSVSALLKIGIRGVTVSDVRGFGAGGSTERQGGSESEDFKVAKVM 62
 QY 63 ETWVKDOVEDVIEKIEEARTGEIGDKIFLLPVSDVIRVTRGERGDKA 112
 DB 62 ENVLPEMDVDIALEALVGNARTEKIDGKIFVSSIEQAIRIRGETGEDA 111
 RESULT 11
 AAB69499
 ID AAB69499 standard; protein; 111 AA.
 XX AAB69499;
 AC AAB69499;
 DT 23-APR-2001 (first entry)
 XX Rhizobium leguminosarum PII protein.
 DE Rhizobium leguminosarum; PII; plant nitrogen regulatory gene; P-II;
 KW nitrogen assimilation; transgenic plant; herbicide screening.
 XX Rhizobium leguminosarum.
 OS Rhizobium leguminosarum.
 XX US6177275-B1.
 PN 23-JAN-2001.
 XX 23-JUL-1997; 97US-00899330.

Db 63 VIISDAQAEVINIIVETARTCKVGDGKVMNIEELVVRTGERGEAA 111
 RESULT 9
 ADA34150
 ID ADA34150 standard; protein; 124 AA.
 XX ADA34150;
 AC ADA34150;
 DT 20-NOV-2003 (first entry)
 XX Acinetobacter baumannii protein #1311.
 DE Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
 KW plant biocontrol agent.
 XX Acinetobacter baumannii.
 OS US6562958-B1.
 PN 13-MAY-2003.
 XX 04-JUN-1999; 99US-00328352.
 XX 09-JUN-1998; 98US-0088701P.
 XX (GENO-) GENOME THERAPEUTICS CORP.
 XX Breton G, Bush D;
 XX WPI; 2003-576092/54.
 XX N-PSDB; ADA34024.
 XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 PT for diagnosing a bacterial disease, as components of antibacterial
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 PT plants.
 XX Example; SEQ ID NO 5437; 328pp; English.
 XX The invention relates to isolated Acinetobacter baumannii nucleic acids.
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents
 CC for diagnosing a bacterial disease, as components of antibacterial
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 CC A. baumannii and other Acinetobacter species in a sample, in screening
 CC compounds for the ability to interfere with the A. baumannii life cycle
 CC or to inhibit A. baumannii infection, and as biocontrol agents for
 CC plants. The present sequence represents the amino acid sequence of an A.
 CC baumannii protein.
 XX SQ Sequence 124 AA;
 Query Match 46.1%; Score 256.5; DB 6; Length 124;
 Best Local Similarity 48.6%; Pred. No. 3.8e-22;
 Matches 53; Conservative 22; Mismatches 33; Indels 1; Gaps 1;
 QY 4 VEAILRPRVSVSALLKIGIRGVTVSDVRGFGAGGSTERQGGSESEDFKVAKVM 63
 DB 16 VTAIVKPKLDDVREALSDIGVQGITVTEVKGFRHKGHTLYRGAERYVD-FLPKVKIE 74
 QY 64 IVWVKDOVEDVIEKIEEARTGEIGDKIFLLPVSDVIRVTRGERGDKA 112
 DB 75 IAISEMVDVIESITRVASTGKIGDKIFVNLQVIRKIRGETGPDA 123
 RESULT 10
 AAB69502
 ID AAB69502 standard; protein; 112 AA.
 XX AAB69502;
 AC AAB69502;
 DT 23-APR-2001 (first entry)

PR 24-JUL-1996; 96US-0022328P.
 XX (UYNV) UN-V NEW YORK STATE.
 PA Coruzzi GM, Lam H, Hsieh M;
 XX WPI; 2001-158572/16.
 DR Novel P-II genes capable of regulating plant nitrogen assimilation,
 XX useful for transgenic plant production, and as probes for isolating
 PT additional genomic clones having P-II gene promoters.
 PT
 XX Example; Fig 1; 35pp; English.
 PS The present sequence is encoded by a nitrogen regulatory PII gene. Novel
 XX plant PII (also called P-II) nucleotide sequences have been isolated.
 CC They are useful for regulating nitrogen assimilation in plants, and in
 CC transgenic plant production. They are also used to engineer organisms
 CC that overexpress wild-type or mutant P-II regulatory proteins. P-II
 CC proteins are useful for in vitro screening of herbicides. P-II
 CC nucleotides may be used as probes for isolating additional genomic clones
 CC with the promoters of P-II genes. P-II promoters are light- and/or
 CC sucrose-inducible, and are suitable for genetic engineering of plants
 CC
 XX Sequence 111 AA;
 SQ
 Query Match 45.1%; Score 251; DB 4; Length 111;
 Best Local Similarity 47.7%; Pred. No. 1.5e-21;
 Matches 51; Conservative 27; Mismatches 27; Indels 2; Gaps 2;
 QY 3 KVEAILRPRVSVQSSALLKIGIRGTVSDVRGFGAGGSTERQGGSEFSEDFVAKVKM 62
 DB 3 KIEAILKPKLDEVRS-LSGVGLQGITVTEAKGFGKQKGTDLVYRGAETVD-FLPKVKY 60
 QY 63 EIVWSKDQVEDVTEKIEIARTGEIGDGKIFLLPVSVDVIRVTGERG 109
 DB 61 EIVLADENAEVTEIRKAAQTGRIGDGKIFVSNVEVIRITGETG 107
 RESULT 12
 AAB69500
 ID AAB69500 standard; protein; 111 AA.
 AC AAB69500;
 XX 23-APR-2001 (first entry)
 DE Bradyrhizobium japonicum PII protein.
 XX Bradyrhizobium japonicum; PII; plant nitrogen regulatory gene; P-II;
 XX nitrogen assimilation; transgenic plant; herbicide screening.
 XX Bradyrhizobium japonicum.
 OS US6177275-B1.
 PN 23-JAN-2001.
 XX 23-JUL-1997; 97US-00899330.
 XX 24-JUL-1996; 96US-0022328P.
 XX (UYNV) UNIV NEW YORK STATE.
 XX Coruzzi GM, Lam H, Hsieh M;
 XX WPI; 2001-158572/16.
 XX Novel P-II genes capable of regulating plant nitrogen assimilation,
 PT useful for transgenic plant production, and as probes for isolating
 PT additional genomic clones having P-II gene promoters.
 XX Example; Fig 1; 35pp; English.

XX The present sequence is encoded by a nitrogen regulatory PII gene. Novel
 CC plant PII (also called P-II) nucleotide sequences have been isolated.
 CC They are useful for regulating nitrogen assimilation in plants, and in
 CC transgenic plant production. They are also used to engineer organisms
 CC that overexpress wild-type or mutant P-II regulatory proteins. P-II
 CC proteins are useful for in vitro screening of herbicides. P-II
 CC nucleotides may be used as probes for isolating additional genomic clones
 CC with the promoters of P-II genes. P-II promoters are light- and/or
 CC sucrose-inducible, and are suitable for genetic engineering of plants
 CC
 XX Sequence 111 AA;
 SQ
 Query Match 44.5%; Score 248; DB 4; Length 111;
 Best Local Similarity 46.7%; Pred. No. 3.4e-21;
 Matches 50; Conservative 26; Mismatches 29; Indels 2; Gaps 2;
 QY 3 KVEAILRPRVSVQSSALLKIGIRGTVSDVRGFGAGGSTERQGGSEFSEDFVAKVKM 62
 DB 3 KIEAILKPKLDEVRS-LSGVGLQGITVTEAKGFGKQKGTDLVYRGAETVD-FLPKVKY 60
 QY 63 EIVWSKDQVEDVTEKIEIARTGEIGDGKIFLLPVSVDVIRVTGERG 109
 DB 61 EIVIGDDLVERAIDAIRRAQTGRIGDGKIFVSNIEAIRITGESG 107
 RESULT 13
 AABG24254
 ID AABG24254 standard; protein; 160 AA.
 XX AABG24254;
 AC AABG24254;
 XX 18-FEB-2002 (first entry)
 DE Novel human diagnostic protein #24245.
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder.
 XX Homo sapiens.
 XX WC200175067-A2.
 PN 11-OCT-2001.
 PD 30-MAR-2001; 2001WO-US008631.
 XX 31-MAR-2000; 2000US-00540217.
 XX 23-AUG-2000; 2000US-00649167.
 XX (HYSE-) HYSEQ INC.
 XX Drmanac RT, Liu C, Tang YT;
 XX MPI; 2001-639362/73.
 XX N-PSDB; AAS88441.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX Claim 20; SEQ ID NO 54613; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food

```

ER 14-JUL-1999; 99DE-01033006.
FR 31-AUG-1999; 99DE-01041378.
FR 31-AUG-1999; 99DE-01041379.
FR 31-AUG-1999; 99DE-01041390.
FR 31-AUG-1999; 99DE-01041391.
FR 03-SEP-1999; 99DE-01042088.
XX
XX (BADI ) BASF AG.
FA
XX
XX Pompejus M, Kroegeer B, Schroeder H, Zelder O, Haberhauer G;
XX
XX WPI: 2001-061974/07.
XX N-PSDB; AAP71296.
XX
XX New isolated Corynebacterium glutamicum nucleic acid for production or
XX modulation of production of fine chemicals such as amino acids,
XX nucleosides, nucleotides, lipide, fatty acids, carbohydrates, vitamins or
XX enzymes.
XX
XX Claim 20; Page 563; 712pp; English.
XX
XX AAF71138 to AAF711357 encode the Corynebacterium glutamicum homeostasis
XX and adaptation (HA) proteins given in AAB79023 to AAB79942. The C.
XX glutamicum HA genes (I) can be used in vectors for expression in host
XX cells and production of fine chemicals, such as, an organic acid,
XX proteinogenic or nonproteinogenic amino acid (preferred), purine or
XX pyrimidine base, nucleoside, nucleotide, lipid, saturated or unsaturated
XX fatty acid, diol, carbohydrate, aromatic compound, vitamin, cofactor,
XX polypeptide or enzyme. The amino acids produced can be lysine, glutamine,
XX glutamate, alanine, aspartate, glycine, serine, threonine, methionine,
XX tyrosine, valine, leucine, isoleucine, arginine, proline, histidine,
XX tyrosine, phenylalanine, or tryptophan. The fine chemical production can
XX be modulated. The presence of (I) or HA proteins encoded by then are used
XX for diagnosing the presence or activity of Corynebacterium diphtheriae.
XX (I) can be used to map the C. glutamicum genome or can be used as markers
XX for genetically engineered Corynebacterium or Brevibacterium. The HA
XX proteins encoded by the (I) are used to maintain homeostasis in C.
XX glutamicum or help the microorganism to adapt to different environmental
XX conditions
XX
XX Sequence 89 AA;
XX
XX Query Match 43.2%; Score 240.5; DB 4; Length 89;
XX Best Local Similarity 49.4%; Pred. No. 2e-20;
XX Matches 44; Conservative 25; Mismatches 19; Indels 1; Gaps 1;
XX
XX QY 24 GIRGVTSVDVRGFGAGGSTERGGSEFSEDFKAVKMEIIVSKQVEDVIEKIEEAR 83
XX Db 1 GVGQMTVTITGGFGQGGKHTEVTRGAEYAVD-FVPKVIIEVIIISDAQAEVINIIVETAR 59
XX
XX QY 84 TGEIGDGKIFLLPVSIVRVRTGERGDKA 112
XX Db 112
XX 60 TSKVGVGDKVMTNIEELVVRVTRGERGEAA 88
XX
XX RESULT 15
XX ID ABB54937
XX ID ABB54937 standard; protein; 112 AA.
XX
XX AC ABB54937;
XX
XX 29-AUG-2003 (revised)
XX 16-MAY-2002 (first entry)
XX
XX Lactococcus lactis protein glnB.
XX
XX Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.
XX
XX Lactococcus lactis; IL1403.
XX
XX FR2807446-A1.
XX
XX 12-OCT-2001.
XX
XX

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 19:10:08 ; Search time 14 Seconds
(without alignments)
776.403 Million cell updates/sec

Title: US-09-756-541-2
Perfect score: 557
Sequence: 1 FYKVEAILRPWRVQVSSALL.....LLPVSVDVIRVTGERGDKAE 113

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	509	91.4	196	D85024	P II nitrogen sens
2	315.5	56.6	112	AH2095	nitrogen regulator
3	314.5	56.5	112	F70310	nitrogen regulator
4	303.5	54.5	112	H81961	nitrogen regulator
5	302.5	54.3	112	B81019	nitrogen regulator
6	300.5	53.9	112	A39696	nitrogen regulator
7	281.5	50.5	112	S13078	nitrogen regulator
8	278.5	50.0	112	S73175	nitrogen regulator
9	278.5	50.0	112	AF0354	nitrogen regulator
10	277.5	49.8	114	D82102	nitrogen regulator
11	276.5	49.6	112	S04377	nitrogen regulator
12	274.5	49.3	112	RGSCP2	nitrogen regulator
13	274.5	49.3	112	C91056	hypothetical prote
14	274.5	49.3	112	G85900	hypothetical prote
15	274.5	49.3	112	AH0826	nitrogen regulator
16	270.5	48.6	112	S52328	nitrogen regulator
17	269.5	48.4	112	S33180	nitrogen regulator
18	268.5	48.2	112	F64062	nitrogen regulator
19	268.5	48.2	112	A82794	nitrogen regulator
20	268.5	48.2	157	A97573	nitrogen regulator
21	267.5	48.0	112	C87493	nitrogen regulator
22	266.5	47.8	112	AD3374	nitrogen regulator
23	265.5	47.7	121	B82302	nitrogen regulator
24	264.5	47.5	112	C64307	probable nitrogen
25	260.5	46.8	112	G64467	nitrogen regulator
26	260.5	46.8	112	G70747	probable nitrogen
27	253.5	45.5	112	D82985	nitrogen regulator
28	251.5	45.2	112	B69468	nitrogen regulator
29	251.5	45.2	112	T35668	nitrogen regulator

30	251	45.1	111	2	B26567	nitrogen regulator
31	249.5	44.8	115	2	D69188	nitrogen regulator
32	248.5	44.6	112	2	H82631	nitrogen regulator
33	248.5	44.6	112	2	AB0560	nitrogen regulator
34	248	44.5	111	2	A33600	nitrogen regulator
35	246.5	44.3	85	2	S76404	nitrogen regulator
36	246.5	44.3	112	2	AG0381	nitrogen regulator
37	245.5	44.1	112	2	B64775	nitrogen regulator
38	245.5	44.1	112	2	H90691	nitrogen regulator
39	245.5	44.1	112	2	D85542	nitrogen regulator
40	244.5	43.9	116	2	AD2915	nitrogen regulator
41	244.5	43.9	116	2	G97689	glnK protein (AJ00
42	243.5	43.7	115	2	B69188	nitrogen regulator
43	237.5	42.6	112	2	D87415	nitrogen regulator
44	234.5	42.1	114	2	H87313	nitrogen regulator
45	218	39.1	113	2	G86823	nitrogen regulator

ALIGNMENTS

RESULT 1

D85024

P II nitrogen sensing protein GLB I [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001

C:Accession: D85024

R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sprin

Nature 402, 769-777, 1999

A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A:Reference number: A85001; MUID:20083488; PMID:10617198

A:Accession: D85024

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-196 <STO>

A:Cross-references: GB:NC_001268; NID:g7268574; PID:NCAB80683.1; GSPDB:GN00140

C:Genetics:

A:Gene: AT4g01900

A:Map position: 4

Query Match 91.4%; Score 509; DB 2; Length 196;
Best Local Similarity 90.3%; Pred. No. 6.2e-40;
Matches 102; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 FYKVEAILRPWRVQVSSALLKIGIRGVTVSDVRGFGAQQGSTERQGGSEFSDKFKVAKV 60
DB 74 FYKVEAILRPWRVQVSSALLKIGIRGVTVSDVRGFGAQQGSTERQGGSEFSDKFKVAKV 133

QY 61 KMEIVVSKQVEDVIEKITEARTGEIGDKIFLLPVSVDVIRVTGERGDKAE 113

DB 134 KMEIVVSKQVEDVIEKITEARTGEIGDKIFLLPVSVDVIRVTGERGDKAE 186

RESULT 2

AH2095

nitrogen regulatory protein P-II glbN [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002

C:Accession: AH2095

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AH2095

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-112 <KOR>

A:Cross-references: GB:BA000019; PID:BA074018.1; PID:gl7131411; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: glbN

C:Gene: GlnB; NMA0447
C:Superfamily: regulatory protein P-II
C:keywords: phosphoprotein
F:51/Binding site: UMP (Tvr) (covalent) #status predicted

A₁Note: the authors translated the codon CCT for residue 83 as Thr.
R.Li, N.; Warren, P.V.; Golbeck, J.H.; Frank, G.; Zuber, H.; Bryant, D.A.
Biochim. Biophys. Acta 1059, 215-225, 1991
A₂Title: Polypeptide composition of the Photosystem I complex and the Photosystem I core

A:Reference number: A56817; MUID:91355213; PMID:1653017

A:Accession: F56817

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-40

A:Experimental source: PCC 6301

C:Genetics:

A:Gene: glnB

C:Function:

A:Description: involved in signal transduction

A:Note: modification state depends on the nitrogen source and the spectral light quality

C:Superfamily: regulatory protein P-II

C:Keywords: phosphoprotein; signal transduction

F:49/Binding site: phosphate (Ser) (covalent) #status predicted

F:51/Binding site: UMP (Tyr) (covalent) #status predicted

Query Match 53.9%; Score 300.5; DB 2; Length 112;
Best Local Similarity 55.5%; Pred. No. 6.7e-21;
Matches 61; Conservative 21; Mismatches 27; Indels 1; Gaps 1;

QY 3 KVEAILRPMRVSVSSALLKIGIRGVTVSDVRGFGAQQGSTERQGGSEFSEDKFVAKVKM 62

DB 3 KIEAILRPMKLDVXIALVNAIGVMTVSEVAGFGKQGTERTYRGSEYTVF-FLPKLKL 61

QY 63 EIVVSKDQVEDVTEKIEEARTEIGDGKIFLLPVSDVIRVVTGERGDKA 112

DB 62 EIVVEDAQVTDVTKIVAAARPGRIGDGKIFVSPDQTIIRIRTGKNADA 111

RESULT 7

S13078

nitrogen regulatory protein P-II - Azospirillum brasilense

C:Species: Azospirillum brasilense

C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 26-Aug-1999

C:Accession: S13078

R:de Zamaroczy, M.; Delorme, F.; Elmerich, C.

Mol. Gen. Genet. 224, 421-430, 1990

A:Title: Characterization of three different nitrogen-regulated promoter regions for the

A:Reference number: S13078; MUID:91094780; PMID:1702507

A:Accession: S13078

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-112 <ZM>

A:Cross-references: EMBL:X51499; NID:g38664; PIDN:CAA35867.1; PID:g38665

A:Experimental source: strain SP7

C:Genetics:

A:Gene: glnB

C:Function:

A:Description: P-II protein synthesis is increased under conditions of nitrogen fixation

C:Superfamily: regulatory protein P-II

C:Keywords: phosphoprotein; signal transduction

F:51/Binding site: UMP (Tyr) (covalent) #status predicted

Query Match 50.5%; Score 281.5; DB 2; Length 112;
Best Local Similarity 49.1%; Pred. No. 3.8e-19;
Matches 54; Conservative 26; Mismatches 29; Indels 1; Gaps 1;

QY 3 KVEAILRPMRVSVSSALLKIGIRGVTVSDVRGFGAQQGSTERQGGSEFSEDKFVAKVKM 62

DB 3 KIEAILRPMKLDVXIALVNAIGVMTVSEVAGFGKQGTERTYRGSEYTVF-FLPKVKI 61

QY 63 EIVVSKDQVEDVTEKIEEARTEIGDGKIFLLPVSDVIRVVTGERGDKA 112

DB 62 EIVVEDSIVERAIEAQAAHTIGRIGDGKIFVTPVBEVVRIRVTGERGKGA 111

RESULT 8

S73175

nitrogen regulatory protein P-II - red alga (Porphyra purpurea) chloroplast

C:Species: chloroplast Porphyra purpurea

C:Date: 19-Mar-1997 #sequence_revision 09-May-1997 #text_change 26-Aug-1999

C:Accession: S73175

R:Reith, M.; Munnholland, J.

Plant Mol. Biol. Rep. 13, 333-335, 1995

A:Title: Complete nucleotide sequence of the Porphyra purpurea chloroplast genome.

A:Reference number: S73108

A:Accession: S73175

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-112 <REI>

A:Cross-references: EMBL:U38804; NID:g1276652; PIDN:AAC08140.1; PID:g1276720

A:Experimental source: strain Avonport

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995

C:Genetics:

A:Gene: glnB

A:Genome: chloroplast

C:Superfamily: regulatory protein P-II

C:Keywords: chloroplast; phosphoprotein; signal transduction

F:51/Binding site: UMP (Tyr) (covalent) #status predicted

Query Match 50.0%; Score 278.5; DB 2; Length 112;
Best Local Similarity 52.7%; Pred. No. 7.2e-19;
Matches 58; Conservative 21; Mismatches 30; Indels 1; Gaps 1;

QY 3 KVEAILRPMRVSVSSALLKIGIRGVTVSDVRGFGAQQGSTERQGGSEFSEDKFVAKVKM 62

DB 3 KIEAILRPMKLDVXIALVNAIGVMTVSEVAGFGKQGTERTYRGSEYTVF-FLPKVKI 61

QY 63 EIVVSKDQVEDVTEKIEEARTEIGDGKIFLLPVSDVIRVVTGERGDKA 112

DB 62 EIVVSDKVNSTIEIIITAKTGEIGDGKIFISDVVEQVIRIRINDLNSAA 111

RESULT 9

AF0354

nitrogen regulatory protein P-II glnB [imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis

C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001

C:Accession: AF0354

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.

deno-Farraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,

Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AF0354

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-112 <XOR>

A:Cross-references: GB:AL590842; PIDN:CAC92161.1; PID:g15980875; GSPDB:GN00175

C:Genetics:

A:Gene: glnB

C:Superfamily: regulatory protein P-II

Query Match 50.0%; Score 278.5; DB 2; Length 112;
Best Local Similarity 49.1%; Pred. No. 7.2e-19;
Matches 54; Conservative 29; Mismatches 26; Indels 1; Gaps 1;

QY 3 KVEAILRPMRVSVSSALLKIGIRGVTVSDVRGFGAQQGSTERQGGSEFSEDKFVAKVKM 62

DB 3 KIDAILRPMKLDVRRALAEVGTMTVTVEKGFGRQKGTETLYRGAEYVVD-FLPKVKI 61

QY 63 EIVVSKDQVEDVTEKIEEARTEIGDGKIFLLPVSDVIRVVTGERGDKA 112

DB 62 EIVVADDIVTCVEALMQTAHTKIGDGKIFVDFVSRVVRIRVTGERQDEEA 111

RESULT 10

D82102

nitrogen regulatory protein P-II [similarity] - Vibrio cholerae (strain N16961 serogroup

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 24-Aug-2001

C:Accession: D82102

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P

l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: D82102
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-114 <HEI>
A:Cross-references: GB:AE004295; GB:AE003852; MID:g9656789; PIDN:AAF95383.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC2239
A:Map position: 1
C:Superfamily: regulatory protein P-II
C:Keywords: phosphoprotein
F:53/Binding site: UMP (Tyr) (covalent) #status predicted

Query Match 49.8%; Score 277.5; DB 2; Length 114;
Best Local Similarity 50.0%; Pred. No. 9.1e-19;
Matches 55; Conservative 25; Mismatches 29; Indels 1; Gaps 1;

QY 3 KVEAILRPRVSVSSALLKIGIRGTVSDVRGFGAQQGSTERQGGSEFSEDKFVAKVKM 62
DB 5 KIEAILKPKLDDVRELAERVGITGTVSEVKFGKQKCHTELYRGAEMVD-FLPKVKL 63
QY 63 EIVVSKQVEDVIEKIEEARTGEIGDKIFLLPVSIVIRVTRGERGDKA 112
DB 64 EIVVTDVADVCVTIETATQTKIGDKGFIITNVERVIRITGEDEDA 113

RESULT 11
S04377
C:Species: *Klebsiella pneumoniae*
A:Title: Identification of the *Klebsiella pneumoniae* glnB gene: nucleotide sequence of
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 26-Aug-1999
C:Accession: S04377
R:Holte, A.; Merrick, M.
Mol. Gen. Genet. 215, 134-138, 1988
A:Title: Identification of the *Klebsiella pneumoniae* glnB gene: nucleotide sequence of
A:Reference number: S04376; MUID:89201233; PMID:2907369
A:Accession: S04377
A:Molecule type: DNA
A:Residues: 1-112 <HOL>
A:Cross-references: EMBL:X14012; MID:g43804; PIDN:CAA32177.1; PID:g43806
A:Experimental source: strain M5al
C:Genetics:
A:Gene: glnB
C:Function:
A:Description: important for the control of glutamine synthase; in nitrogen-limiting con-
to form P-II-UMP; P-II-UMP allows the deadenylation of glutamine synthetase, thus activa-
GS
A:Note: P-II indirectly controls the transcription of the GS gene (glnA) by preventing N
ted, these events are reversed
C:Superfamily: regulatory protein P-II
C:Keywords: phosphoprotein; signal transduction
F:53/Binding site: UMP (Tyr) (covalent) #status predicted

Query Match 49.6%; Score 276.5; DB 2; Length 112;
Best Local Similarity 50.0%; Pred. No. 1.1e-18;
Matches 55; Conservative 26; Mismatches 28; Indels 1; Gaps 1;

QY 3 KVEAILRPRVSVSSALLKIGIRGTVSDVRGFGAQQGSTERQGGSEFSEDKFVAKVKM 62
DB 3 KIDAILKPKLDDVRELAERVGITGTVSEVKFGKQKCHTELYRGAEMVD-FLPKVKI 61
QY 63 EIVVSKQVEDVIEKIEEARTGEIGDKIFLLPVSIVIRVTRGERGDKA 112
DB 62 EIVVTDVADVCVTIETATQTKIGDKGFIITNVERVIRITGEDEDA 111

RESULT 12
RGECP2
C:Species: *Escherichia coli*
A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: D82102
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-114 <HEI>
A:Cross-references: GB:AE004295; GB:AE003852; MID:g9656789; PIDN:AAF95383.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC2239
A:Map position: 1
C:Superfamily: regulatory protein P-II
C:Keywords: phosphoprotein
F:53/Binding site: UMP (Tyr) (covalent) #status predicted

C:Date: 30-Jun-1988 #sequence_revision 10-Nov-1995 #text_change 01-Mar-2002
C:Accession: C49940; S15991; A29307; S37753; H65032; S31361
R:Liu, J.; Magasanik, B.
J. Bacteriol. 175, 7441-7449, 1993
A:Title: The glnB region of the *Escherichia coli* chromosome.
A:Reference number: A49940; MUID:94042920; PMID:8226691
A:Accession: C49940
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-112 <LIU>
A:Cross-references: GB:S67014; MID:g455660; PIDN:AAB28779.1; PID:g455663
A:Note: sequence extracted from NCBI backbone (NCBIN:139878, NCBIP:139882)
R:Vasudevan, S.G.; Armarego, W.L.F.; Shaw, D.C.; Lilley, P.E.; Dixon, N.E.; Poole, R.K.
Mol. Gen. Genet. 226, 49-58, 1991
A:Title: Isolation and nucleotide sequence of the hmp gene that encodes a haemoglobin-l
A:Reference number: S15991; MUID:91238719; PMID:2034230
A:Accession: S15991
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-112 <VAS>
A:Cross-references: GB:X58872; MID:g41730; PIDN:CAA41683.1; PID:g41732
A:Experimental source: strain K-12
R:Son, H.S.; Rhee, S.G.
J. Biol. Chem. 262, 8690-8695, 1987
A:Title: Cascade control of *Escherichia coli* glutamine synthetase: purification and pro
A:Reference number: A29307; MUID:87250488; PMID:2885322
A:Accession: A29307
A:Molecule type: DNA
A:Residues: 1-18,'R',20-80,'E',83-102,'VP' <SON>
A:Cross-references: GB:M16778; MID:g146165; PIDN:AAA23883.1; PID:g146166
R:van Heeswijk, W.C.; Rabenberg, M.; Westerhoff, H.V.; Kahn, D.
Mol. Microbiol. 9, 443-457, 1993
A:Title: The genes of the glutamine synthetase adenyllylation cascade are not regulated i
A:Reference number: S36254; MUID:94018640; PMID:8412694
A:Accession: S37753
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-12 <VAN>
A:Cross-references: EMBL:Z21843; MID:g49395; PIDN:CAA79890.1; PID:g49397
A:Experimental source: strain K-12, substrain W310
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
A:Rose, D.J.; Mau, B.; Shaoh, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of *Escherichia coli* K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: H65032
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-112 <BLAT>
A:Cross-references: GB:AE000341; GB:U00096; MID:g1788899; PIDN:AAC75606.1; PID:g1788904
A:Experimental source: strain K-12, substrain MGL655
C:Genetics:
A:Gene: glnB
C:Function: <GEN>
A:Pathway: nitrogen regulation pathway
A:Note: this protein participates indirectly in activation of glutamine synthase activi
A:Note: under nitrogen-limiting conditions it is covalently uridylylated; in nitrogen e
transferase / uridylyl-removing enzyme which is regulated by the glutamine/alpha-ketogl
C:Function: <NTRB>
A:Description: de-uridylylated P-II forms a complex with nitrogen regulation protein II
the uridylylated form of P-II does not complex with ntrB; free ntrB phosphorylates ntr
A:Note: phosphorylated nitrogen regulation protein I (ntrC) activates transcription of I
C:Function: <NTR>
A:Description: uridylylated P-II forms a complex with adenyllyltransferase; this complex
yllyltransferase deadenylates glutamine synthase
A:Note: free glutamine synthase is active; adenyllylated glutamine synthase is inactive
C:Superfamily: regulatory protein P-II
C:Keywords: phosphoprotein; signal transduction
F:53/Binding site: UMP (Tyr) (covalent) #status experimental

Query Match 49.3%; Score 274.5; DB 1; Length 112;
Best Local Similarity 50.0%; Pred. No. 1.7e-18;
Matches 55; Conservative 25; Mismatches 29; Indels 1; Gaps 1;

QY 63 EIVVSKQVEDVIEKILIEARTGETCDGKIQLLPUSDVIRVRTGERGDKA 112
||||| : : ||||| : : ||||| : : ||||| : : ||||| : :
Db 62 EIVVPDDIVDTCTVTIINTAQTGKTGGDKIFVDFVARVIRIRTGEDDAA 111

RESULT 15
AH0826
nitrogen regulatory protein p-II STY2808 [imported] - Salmonella enterica subsp. enteric
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: This species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AH0826
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUIDB:21534947; PMID:11677608
A;Accession: AH0826
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-112 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD02764.1; PID:g16503774; GSPDB:GN00176
C;Genetics:
A;Gene: STY2808
C;Superfamily: regulatory protein p-II

Query Match	49.3%	Score 274.5;	DB 2;	Length 112;
Best Local Similarity	50.0%;	Pred. No. 1.7e-18;		
Matches	55; Conservative	25; Mismatches	29; Indels	1; Gaps 1
Qy	3	KVRAILRPMRVQSVALLKIGIRGVTVSDVRGFAQGSGSTERGGSEFSEDKFWAKVM	62	
	:::::>:::	:::::>:::	:::::>:::	:::::>:::
Dd	3	KIDAILKPFLDVRALAEVGITGMVTIEWGFGRKGHTELRYGAEMVD-FLPKVKI	61	
	:::::>:::	:::::>:::	:::::>:::	:::::>:::
Qy	63	EIVWSKDQEDVTEKIIEBATGEIGDKJELLFPVSQVIRVTRGERDKA	112	
	:::::>:::	:::::>:::	:::::>:::	:::::>:::
Dd	62	EIVVPDDIVDTCTTIIRTAQTCKIGDKJFFVFDAVRVIRTGTGEEDAA	111	
	:::::::~:	:::::>:::	:::::>:::	:::::>:::

Search completed: May 24, 2004, 19:15:40
Job time : 14 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 19:06:17 ; Search time 10.5 seconds
(without alignments)
560.374 Million cell updates/sec

Title: US-09-756-541-2

Perfect score: 557

Sequence: 1 FYKVEAILRPWRVQVSSAL.....LLPVSDVIRVTRGGRKAE 113

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	315.5	56.6	112	1	GLNB_NOSPU
2	314.5	56.5	112	1	GLNB_AQARE
3	314.5	56.5	112	1	GLNB_FREDI
4	311.5	55.9	112	1	GLNB_SYNP3
5	306.5	55.0	112	1	GLNB_SYNP7
6	286.5	51.4	112	1	GLNB_PASMU
7	281.5	50.5	112	1	GLNB_AZOB
8	278.5	50.0	112	1	GLNB_PORPU
9	276.5	49.6	112	1	GLNB_KLEOX
10	274.5	49.3	112	1	GLNB_ECOLI
11	270.5	48.6	112	1	GLNB_RHORU
12	269.5	48.4	112	1	GLNB_RHIME
13	269.5	48.4	112	1	GLNB_RHOSH
14	268.5	48.2	112	1	GLNB_HAEIN
15	267.5	48.0	112	1	GLNB_RHILLO
16	264.5	47.5	112	1	Y059_METJA
17	260.5	46.8	112	1	GLNB_BRAJA
18	260.5	46.8	112	1	GLNB_MYCTU
19	260.5	46.8	112	1	YD44_METJA
20	255.5	45.9	112	1	GLNB_RHOCA
21	254.5	45.7	113	1	GLNB_CVACA
22	251	45.1	111	1	GLNB_RHLIV
23	249.5	44.8	115	1	GLNB_METTH
24	245.5	44.1	112	1	GLNB_ECOLI
25	243.5	43.7	115	1	GLNB_RHLET
26	232.5	41.7	112	1	GLNB_RHLET
27	195.5	35.1	116	1	NRGB_BACSU
28	181.5	32.6	123	1	GLN2_METBA
29	170.5	30.6	125	1	GLN4_METBA
30	167.5	30.1	128	1	GLN2_METTL
31	166	29.8	121	1	GLN2_METMP
32	165	29.6	122	1	GLN2_METIV
33	163.5	29.4	105	1	GLN1_METIV

34	159	28.5	121	1	GLN2_METTM	Q50787 methanobact
35	153.5	27.6	105	1	GLN1_METTL	P25771 methanococc
36	149.5	26.8	105	1	GLN1_METTM	Q50786 methanobact
37	144.5	25.9	105	1	GLN1_METBA	P54808 methanobact
38	140.5	25.2	105	1	GLN1_METMP	P71524 methanococc
39	135.5	24.3	105	1	GLN3_METBA	P54807 methanobact
40	79.5	14.3	307	1	GSBB_PROMP	Q07989 prochloroco
41	75.5	13.6	305	1	PARI_HELAS	Q07981 helix asper
42	75	13.5	187	1	KIP2_MOUSE	Q92309 mus musculu
43	72.5	13.0	529	1	IMDH_MYCTU	Q50715 mycobacteri
44	72	12.9	187	1	KIP2_HUMAN	O75838 homo sapien
45	70.5	12.7	573	1	CH60_HUMAN	P10809 homo sapien

ALIGNMENTS

RESULT 1

ID	GLNB_NOSPU	STANDARD;	PRT;	112 AA.
AC	Q30794;			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Nitrogen regulatory protein P-II (PII signal transducing protein).			
GN	GLNB.			
OS	Nostoc punctiforme.			
OC	Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.			
OX	NCBI_TaxID=63737;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 29133 / PCC 73102;			
RX	MEDLINE=98304077; Pubmed=9639924;			
RA	Hanson T.E., Forchhammer K., Tandeau de Marsac N., Meeks J.C.;			
RT	*Characterization of the glnB gene product of Nostoc punctiforme			
RT	strain ATCC 29133: glnB or the PII protein may be essential.";			
RL	Microbiology 144:1537-1547(1998).			
CC	-I- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE GS			
CC	GENE (GLNA). P-II PREVENTS NR-II CATALYZED CONVERSION OF NR-I TO			
CC	NR-I-PHOSPHATE, THE TRANSCRIPTION ACTIVATOR OF GLNA. WHEN P-II			
CC	IS PHOSPHORYLATED, THESE EVENTS ARE REVERSED (BY SIMILARITY).			
CC	-I- FUNCTION: IN NITROGEN-LIMITING CONDITIONS, WHEN THE RATIO OF GLN			
CC	TO 2-KETOGLUTARATE DECREASES, P-II IS PHOSPHORYLATED WHICH ALLOWS			
CC	THE DEADENYLATION OF GLUTAMINE SYNTHETASE (GS), THUS ACTIVATING			
CC	THE ENZYME (BY SIMILARITY).			
CC	-I- SUBUNIT: Homotrimer (by similarity).			
CC	-I- PTM: PHOSPHORYLATION DEPENDENT ON THE NITROGEN SOURCE AND SPECTRAL			
CC	LIGHT QUALITY (BY SIMILARITY).			
CC	-I- SIMILARITY: Belongs to the P(II) protein family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement [See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch].			
CC	-----			
CC	EMBL; AF017419; AAC26348.1; .			
DR	HSSP; P05826; 2P1I.			
DR	InterPro; IPR002187; PII_glnB			
DR	InterPro; IPR002332; PII_GlnB_UMP_S.			
DR	Pfam; PF00543; P-II; 1.			
DR	PRINTS; PR00340; PIIGLNB.			
DR	ProDom; PD001194; PII_glnB; 1.			
DR	PROSITE; PS00496; PII_GlnB_UMP; 1.			
DR	PROSITE; PS00638; PII_GlnB_CTER; 1.			
DR	Transcription regulation; Nitrogen fixation; Phosphorylation.			
FW	MOD RES 49 49 PHOSPHORYLATION (PROBABLE).			
FT	SEQUENCE 112 AA; 12478 MW; 9C224C38B67583A CRC64;			
SQ				

Query Match

Best Local Similarity 56.6%; Score 315.5; DB 1; Length 112;

Pred. No. 1e-22;

Matches 64; Conservative 22; Mismatches 23; Indels 1; Gaps 1;

QY 3 KVEAIRLRWRSVSSALLKIGIRGTVSDVVRGFGAQQGSTERQGGSEFSEDFKFAVKVM 62
 Db 3 KVEAIRLRWRSVSSALLKIGIRGTVSDVVRGFGAQQGSTERQGGSEFSEDFKFAVKVM 62
 QY 63 EIVVSKQDVEDVTEKIEEARTGEIGDGKIFLLPVSDVIRVIRGTCRDKA 112
 Db 62 EIVVSDNQVDMVVDKIIAAARTGEIGDGKIFISFVEQVIRIRTGKNTTEA 111

RESULT 2
 GLOBE AQAUE
 ID GLOBE AQAUE STANDARD; PRT; 112 AA.
 AC 066513;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Nitrogen regulatory protein P-II.
 GN GLOBE OR AQ 109.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
 OX NCBI_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VF5;
 RX MEDLINE=98196666; PubMed=9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 aeolicus";
 RL Nature 392:353-358(1998).
 CC -!- FUNCTION: IN NITROGEN-LIMITING CONDITIONS, WHEN THE RATIO OF GLN
 TO 2-KETOGLUTARATE DECREASES, P-II IS URIDYLATED TO P-II-UMP. P-
 CC 11-UMP ALLOWS THE DEADENYLATION OF GLUTAMINE SYNTHETASE (GS).
 CC THIS ACTIVATING AND PROMOTES THE ADENYLATION OF GS (BY
 CC IS DEURIDYLATED AND PROMOTES THE ADENYLATION OF GS (BY
 CC SIMILARITY).
 CC -!- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE GS
 CC GENE (GLNA). P-II PREVENTS NR-II CATALYZED CONVERSION OF NR-I TO
 CC NR-I-PHOSPHATE, THE TRANSCRIPTIONAL ACTIVATOR OF GLNA. WHEN P-II
 CC IS URIDYLATED TO P-II-UMP, THESE EVENTS ARE REVERSED (BY
 CC SIMILARITY).
 CC -!- SUBUNIT: Homotrimer (By similarity).
 CC -!- SIMILARITY: Belongs to the P(II) protein family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).

EMBL; AE000674; AAC06473.1; -
 DR PIR; F70310; F70310.
 DR HSSP; P05826; 2PII.
 DR InterPro; IPR002187; PII_glnB.
 DR PRINTS; PR00340; PII_glnB.
 DR PRODOM; PD001194; PII_glnB; 1.
 DR PROSITE; PS00496; PII_GlnB_UMP; 1.
 DR PROSITE; PS00638; PII_GlnB_CTER; 1.
 KW Transcription regulation; Nitrogen fixation; Complete proteome.
 FT BINDING 51 UMP (BY SIMILARITY).
 SQ SEQUENCE 112 AA; 12497 MW; 0E44B4B171A6233B CRC64;

Query Match 56.5%; Score 314.5; DB 1; Length 112;
 Best Local Similarity 53.6%; Pred. No. 1.3e-22;
 Matches 59; Conservative 29; Mismatches 21; Indels 1; Gaps 1;

QY 3 KVEAIRLRWRSVSSALLKIGIRGTVSDVVRGFGAQQGSTERQGGSEFSEDFKFAVKVM 62
 Db 3 KVEAIRLRWRSVSSALLKIGIRGTVSDVVRGFGAQQGSTERQGGSEFSEDFKFAVKVM 61
 QY 63 EIVVSKQDVEDVTEKIEEARTGEIGDGKIFLLPVSDVIRVIRGTCRDKA 112
 Db 62 EIVVSDNQVDMVVDKIIAAARTGEIGDGKIFISFVEQVIRIRTGKNTTEA 111

RESULT 3
 GLOBE FREDI
 ID GLOBE FREDI STANDARD; PRT; 112 AA.
 AC 047894;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Nitrogen regulatory protein P-II (PII signal transducing protein).
 GN GLOBE.
 OS Fremyella diplosiphon (Calothrix PCC 7601).
 OC Bacteria; Cyanobacteria; Nostocales; Microchaetaceae; Fremyella.
 OX NCBI_TaxID=1197;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Liotenberg S., Castets A.M., Campbell P., Hounard J.,
 RA Tandeau de Marsac N.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE GS
 CC GENE (GLNA). P-II PREVENTS NR-II CATALYZED CONVERSION OF NR-I TO
 CC NR-I-PHOSPHATE, THE TRANSCRIPTIONAL ACTIVATOR OF GLNA. WHEN P-II
 CC IS PHOSPHORYLATED, THESE EVENTS ARE REVERSED (BY SIMILARITY). GLN
 CC -!- FUNCTION: IN NITROGEN-LIMITING CONDITIONS, WHEN THE RATIO OF GLN
 CC TO 2-KETOGLUTARATE DECREASES, P-II IS PHOSPHORYLATED WHICH ALLOWS
 CC THE DEADENYLATION OF GLUTAMINE SYNTHETASE (GS), THUS ACTIVATING
 CC THE ENZYME (BY SIMILARITY).
 CC -!- SUBUNIT: Homotrimer (By similarity).
 CC -!- PTM: PHOSPHORYLATION DEPENDENT ON THE NITROGEN SOURCE AND SPECTRAL
 CC LIGHT QUALITY (BY SIMILARITY).
 CC -!- SIMILARITY: Belongs to the P(II) protein family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).

EMBL; X97327; CAA65992.1; -
 DR HSSP; P05826; 2PII.
 DR InterPro; IPR002187; PII_glnB.
 DR PRINTS; PR00340; P-II; 1.
 DR PRODOM; PD001194; PII_glnB; 1.
 DR PROSITE; PS00496; PII_GlnB_UMP; 1.
 DR PROSITE; PS00638; PII_GlnB_CTER; 1.
 KW Transcription regulation; Nitrogen fixation; Phosphorylation.
 FT MOD_RES 49 49 PHOSPHORYLATION (PROBABLE).
 SQ SEQUENCE 112 AA; 12478 MW; F7A854C393C4596B CRC64;

Query Match 56.5%; Score 314.5; DB 1; Length 112;
 Best Local Similarity 57.3%; Pred. No. 1.3e-22;
 Matches 63; Conservative 23; Mismatches 23; Indels 1; Gaps 1;

QY 3 KVEAIRLRWRSVSSALLKIGIRGTVSDVVRGFGAQQGSTERQGGSEFSEDFKFAVKVM 62
 Db 3 KVEAIRLRWRSVSSALLKIGIRGTVSDVVRGFGAQQGSTERQGGSEFSEDFKFAVKVM 61
 QY 63 EIVVSKQDVEDVTEKIEEARTGEIGDGKIFLLPVSDVIRVIRGTCRDKA 112
 Db 62 EIVVSDNQVDMVVDKIIAAARTGEIGDGKIFISFVEQVIRIRTGKNTTEA 111

RESULT 4
 Glnb_Syn3
 ID Glnb_Syn3 STANDARD; PRT; 112 AA.
 AC Q55247;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Nitrogen regulatory protein P-II (PII signal transducing protein).
 GN Glnb OR SSI0707.
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98088000; PubMed=9426594;
 RA Garcia-Dominguez M., Florencio F.J.;
 RT "Nitrogen availability and electron transport control the expression
 of glbB gene (encoding PII protein) in the cyanobacterium
 Synechocystis sp. PCC 6803.";
 RL Plant Mol. Biol. 35:723-734 (1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hiroseawa M., Sugiyama M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
 RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 RA Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 Synechocystis sp. strain PCC6803. II. Sequence determination of the
 entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136 (1996).
 CC -!- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE GS
 CC GENE (GLNA). P-II PREVENTS NR-II CATALYZED CONVERSION OF NR-I TO
 CC NR-I-PHOSPHATE. THE TRANSCRIPTIONAL ACTIVATOR OF GLNA, WHEN P-II
 CC IS PHOSPHORYLATED, THESE EVENTS ARE REVERSED (BY SIMILARITY).
 CC -!- FUNCTION: IN NITROGEN-LIMITING CONDITIONS, WHEN THE RATIO OF GLN
 CC TO 2-KETOGLUTARATE DECREASES, P-II IS PHOSPHORYLATED WHICH ALLOWS
 CC THE DEADENYLATION OF GLUTAMINE SYNTHETASE (GS), THUS ACTIVATING
 CC THE ENZYME (BY SIMILARITY).
 CC -!- SUBUNIT: Homotrimer (by similarity).
 CC -!- PTM: PHOSPHORYLATION DEPENDENT ON THE NITROGEN SOURCE AND SPECTRAL
 CC LIGHT QUALITY (BY SIMILARITY).
 CC -!- SIMILARITY: Belongs to the P(II) protein family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL; X37496; CAA66127.1; --
 DR EMBL; D90915; BAAL8533.1; ALT_INIT.
 DR HSP; P05826; 2P11.
 DR InterPro; IPR002187; PII_glnb.
 DR InterPro; IPR002332; PII_Glnb_UMP_S.
 DR Pfam; PF00543; P-II; 1.
 DR PRINTS; PR00340; PIIGLNB.
 DR ProDom; PD031194; PII_glnb; 1.
 DR PROSITE; PS00496; PII_Glnb_UMP; 1.
 DR PROSITE; PS00638; PII_Glnb_CTER; 1.
 KW Transcription regulation; Nitrogen fixation; Phosphorylation;
 KW Complete proteome.
 FT MOD_RES 49 49 PHOSPHORYLATION (PROBABLE).
 SQ SEQUENCE 112 AA; 12397 MM; F9ABD0F5C173B799 CRC64;
 Query Match 55.9%; Score 311.5; DB 1; Length 112;
 Best Local Similarity 55.5%; Pred. No. 2.5e-22;
 Matches 61; Conservative 25; Mismatches 23; Indels 1; Gaps 1;
 QY 3 KVEAILRPWRVQSSALLKIGINGVTVDVRGFGAOGGSTERGGSEFSEDKFVAKVM 62

Db 3 KVEAILRPFKLDEVKIALVNGIVGTVSEVRGFGKQGTERTYRGSEYVE-FLQKIKI 61
 QY 63 EIVSKQQVEDVIEKIEIARTGEIGDKIFLLPVSDFVIRVTGERGDKA 112
 Db 62 EIWDGQGVDMVVDKLVSAARTGEIGDKIFISPDVSVVIRTKGDKTEA 111
 RESULT 5
 Glnb_Syn7
 ID Glnb_Syn7 STANDARD; PRT; 112 AA.
 AC P80016;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Nitrogen regulatory protein P-II (PII signal transducing protein).
 GN Glnb.
 OS Synechococcus sp. (strain PCC 7942) (Anacystis nidulans R2), and
 OS Synechococcus sp. (strain PCC 6301) (Anacystis nidulans).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
 OX NCBI_TaxID=1140, 1139;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=PCC 7942;
 RX MEDLINE=91271233; PubMed=1905010;
 RA Tainorenas N.F., Castets A.M., Harrison M.A., Allen J.F.,
 RA Tandeau de Marsac N.;
 RT "Photosynthetic electron transport controls nitrogen assimilation in
 cyanobacteria by means of posttranslational modification of the glbB
 gene product.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:4565-4569 (1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=PCC 6301;
 RX MEDLINE=90249505; PubMed=2110911;
 RA Harrison M.A., Keen J.N., Findlay J.B.C., Allen J.F.;
 RT "Modification of a glbB-like gene product by photosynthetic electron
 transport in the cyanobacterium Synechococcus 6301.";
 RL PNAS Lett. 264:25-28 (1990).
 RN [3]
 RP SEQUENCE OF 1-30.
 RX STRAIN=PCC 6301;
 RX MEDLINE=91355213; PubMed=1653017;
 RA Li N., Warren P.V., Golbeck J.H., Frank G., Zuber H., Bryant D.A.;
 RT "Polypeptide composition of the photosystem I complex and the
 Photosystem I core protein from Synechococcus sp. PCC 6301.";
 RL Biochim. Biophys. Acta 1059:215-225 (1991).
 RN [5]
 RP PHOSPHORYLATION.
 RX STRAIN=PCC 7942;
 RX MEDLINE=94110245; PubMed=8282715;
 RA Forchhammer K., Tandeau de Marsac N.;
 RT "The PII protein in the cyanobacterium Synechococcus sp. strain PCC
 7942 is modified by serine phosphorylation and signals the cellular
 N-status.";
 RL J. Bacteriol. 176:84-91 (1994).
 RN [6]
 RP PHOSPHORYLATION.
 RX STRAIN=PCC 7942;
 RX MEDLINE=96011366; PubMed=7592328;
 RA Forchhammer K., Tandeau de Marsac N.;
 RT "Phosphorylation of the PII protein (glbB gene product) in the
 cyanobacterium Synechococcus sp. strain PCC 7942: analysis of in
 vitro kinase activity.";
 RL J. Bacteriol. 177:5812-5817 (1995).
 CC -!- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE GS
 CC GENE (GLNA). P-II PREVENTS NR-II CATALYZED CONVERSION OF NR-I TO

CC NR-1-PHOSPHATE, THE TRANSCRIPTIONAL ACTIVATOR OF GLNA. WHEN P-II
CC IS PHOSPHORYLATED, THESE EVENTS ARE REVERSED.
CC -1- FUNCTION: IN NITROGEN-LIMITING CONDITIONS, WHEN THE RATIO OF GLN
CC TO 2-KETOGLUTARATE DECREASES, P-II IS PHOSPHORYLATED WHICH ALLOWS
CC THE DEADENYLATION OF GLUTAMINE SYNTHETASE (GS), THUS ACTIVATING
CC THE ENZYME.
CC -1- SUBUNIT: Homotrimer.
CC -1- PTM: PHOSPHORYLATION DEPENDENT ON THE NITROGEN SOURCE AND SPECTRAL
CC LIGHT QUALITY.
CC -1- SIMILARITY: Belongs to the P(II) protein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M62447; AAA27312.1; -.
CC EMBL; AF079137; AAF04333.1; -.
CC PIR; A36596; A39696.
CC HSP; P05826; 2PIL.
CC PROSITE; PS0016; -.
CC InterPro; IPR002187; PII_glnB.
CC Pfam; PF00543; P-II; 1.
CC PRINTS; PR00340; PII_GlnB.
CC ProDom; PD001194; PII_glnB; 1.
CC PROSITE; PS00496; PII_GlnB_UMP; 1.
CC PROSITE; PS00638; PII_GlnB_CTR; 1.
CC Transcription regulation; Nitrogen fixation; Phosphorylation.
CC MOD RES 49 49 PHOSPHORYLATION.
CC FT CONFLICT 83 83 T -> P (IN REF. 1; AAA27312).
CC FT SEQUENCE 112 AA; 12391 MW; 5F44B64CBFF3C559 CRC64;
CC
CC Query Match 55.0%; Score 306.5; DB 1; Length 112;
CC Best Local Similarity 56.4%; Pred. No. 7.2e-22;
CC Matches 62; Conservative 21; Mismatches 26; Indels 1; Gaps 1;
CC
CC 3 KVEATLRPRVSVSSALLKIGIRGVTVSDVRGFGAQQGSTERQGGSEFSEDFVAKVKM 62
CC 3 KIEAIRPFKLDVRESLDVGITGVTVSEVGRGFRQKQTERYRGSEYTVB-FLQKLKL 61
CC
CC 63 EIVVSKQDVEDVIEKIIIEARTGEIGDGKIFLLPVSDVIRVIRTGKDKA 112
CC 62 EIVVEDAQVTVDKIVAAARTGEIGDGKIFVSPVDQTIIRIRTKGNADA 111
CC
CC RESULT 6
CC ID GLNB_PASMU STANDARD; PRT; 112 AA.
CC AC Q9CJL1
CC DT 16-OCT-2001 (Rel. 40, Created)
CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Nitrogen regulatory protein P-II.
CC GN GLNB OR PM2004.
CC OS Pasteurella multocida.
CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
CC OC Pasteurellaceae; Pasteurella.
CC OX NCBI_TaxID=747;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=PM70;
CC RX MEDLINE=21145866; PubMed=11248100;
CC RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
CC "Complete genomic sequence of Pasteurella multocida pm70.";
CC Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE
CC GLUTAMINE SYNTHETASE GENE (GLNA). P-II PREVENTS NR-II CATALYZED
CC CONVERSION OF NR-I TO NR-I-PHOSPHATE, THE TRANSCRIPTIONAL
CC ACTIVATOR OF GLNA. WHEN P-II IS URIDYLATED TO P-II-UMP, THESE

CC EVENTS ARE REVERSED. WHEN THE RATIO OF GLN TO 2-KETOGLUTARATE
CC DECREASES, P-II IS URIDYLATED TO P-II-UMP, WHICH CAUSES THE
CC DEADENYLATION OF GLUTAMINE SYNTHETASE BY GLNE, SO ACTIVATING THE
CC ENZYME (BY SIMILARITY).
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- PTM: Uridylylated/deuridylylated by glnd (By similarity).
CC -1- SIMILARITY: Belongs to the P(II) protein family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE006236; AA04088.1; -.
CC HSP; P05826; 2PIL.
CC InterPro; IPR002187; PII_glnB.
CC Pfam; PF00543; P-II; 1.
CC PRINTS; PR00340; PII_GlnB.
CC ProDom; PD001194; PII_glnB; 1.
CC PROSITE; PS00496; PII_GlnB_UMP; 1.
CC PROSITE; PS00638; PII_GlnB_CTR; 1.
CC Transcription regulation; Nitrogen fixation; Complete proteome.
CC FT BINDING 51 51 UMP (BY SIMILARITY).
CC FT SEQUENCE 112 AA; 12684 MW; OCC9D24651F47642 CRC64;
CC
CC Query Match 51.4%; Score 286.5; DB 1; Length 112;
CC Best Local Similarity 52.7%; Pred. No. 5.2e-20;
CC Matches 58; Conservative 23; Mismatches 28; Indels 1; Gaps 1;
CC
CC 3 KVEATLRPRVSVSSALLKIGIRGVTVSDVRGFGAQQGSTERQGGSEFSEDFVAKVKM 62
CC 3 KIEAIRPFKLDVRESLDVGITGVTVSEVGRGFRQKQTERYRGSEYTVB-FLQKLKL 61
CC
CC 63 EIVVSKQDVEDVIEKIIIEARTGEIGDGKIFLLPVSDVIRVIRTGKDKA 112
CC 62 EIVVTEQVDQICIEALMETAQTKIGDGKIFVYDVERVIRINTGEENEDA 111
CC
CC RESULT 7
CC ID GLNB_AZOBR STANDARD; PRT; 112 AA.
CC AC P21193;
CC DT 01-MAY-1991 (Rel. 18, Created)
CC DT 01-MAY-1991 (Rel. 18, Last sequence update)
CC DT 01-NOV-1995 (Rel. 32, Last annotation update)
CC DE Nitrogen regulatory protein P-II.
CC GN GLNB.
CC OS Azospirillum brasilense.
CC OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
CC OC Rhodospirillaceae; Azospirillum.
CC OX NCBI_TaxID=192;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=Sp7;
CC RX MEDLINE=91094780; PubMed=1702507;
CC RA de Zamaroczy M., Belorme F., Elmerich C.;
CC "Characterization of three different nitrogen-regulated promoter
CC regions for the expression of glnB and glnA in Azospirillum
CC brasilense.";
CC RT Mol. Gen. Genet. 224:421-430(1990).
CC RN [2]
CC RP SEQUENCE OF 88-112 FROM N.A.
CC RC STRAIN=Sp7;
CC RX MEDLINE=87076765; PubMed=2878685;
CC RA Bozouklian H., Elmerich C.;
CC "Nucleotide sequence of the Azospirillum brasilense Sp7 glutamine
CC synthetase structural gene.";
CC RL Biochimie 68:1181-1187(1986).
CC -1- FUNCTION: IN NITROGEN-LIMITING CONDITIONS, WHEN THE RATIO OF GLN

TO 2-KETOGLUTARATE DECREASES, P-II IS URIDYLATED TO P-II-UMP.
P-II-UMP ALLOWS THE DEADENYLATION OF GLUTAMINE SYNTHETASE (GS),
THUS ACTIVATING THE ENZYME. CONVERSELY, IN NITROGEN EXCESS P-II
IS URIDYLATED AND PROMOTES THE ADENYLATION OF GS.
-I- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE GS
GENE (GLNA). P-II PREVENTS NR-II CATALYZED CONVERSION OF NR-I TO
NR-I-PHOSPHATE, THE TRANSCRIPTIONAL ACTIVATOR OF GLNA. WHEN P-II
IS URIDYLATED TO P-II-UMP, THESE EVENTS ARE REVERSED.
-I- SUBUNIT: Homotrimer (By similarity).
-I- SIMILARITY: Belongs to the P(II) protein family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL: U38834; AAC08140.1; -
PIR: S73175; S73175.
HSSP: P05826; 2PII.
InterPro: IPR002187; PII_glnB.
InterPro: IPR002332; PII_glnB_UMP_S.
Pfam: PF00543; P-II; 1.
PRINTS: PR00340; PIIGLNB.
ProDom: PD001194; PII_glnB; 1.
PROSITE: PS00496; PII_GLNBL_UMP; 1.
PROSITE: PS00638; PII_GLNBL_CTR; 1.
KW Transcription regulation; Nitrogen fixation; Chloroplast.
FT BINDING 51 UMP (By similarity).
SQ SEQUENCE 112 AA; 12320 MW; 29416AA49FC37A18 CRC64;

Query Match 50.0%; Score 278.5; DB 1; Length 112;
Best Local Similarity 52.7%; Pred. No. 2.9e-19;
Matches 58; Conservative 21; Mismatches 30; Indels 1; Gaps 2;

QY 3 KVEATLRPRVSVSALLKIGIRGVTVSDVRGFGAQQGSTRFGQGSFSEDFKFAKVKM 62
DB 3 KIEAIRPKLDEVEALHEVGIKGVTVGKGRQKGTLYRGAEVVD-FLPKVKI 61

QY 63 EIVSKDQVEDVIEKIEEARTGEIGDKIFLLPVSVDVIRVTRGEGDKA 112
DB 62 EIVSDDKVNSITELIKTKTGTGIGDKGKIFISDVQEVIRITNDLNSAA 111

RESULT 9
GLNB_PORPU STANDARD; PRT; 112 AA.
ID GLNB_PORPU STANDARD; PRT; 112 AA.
AC P11254; 1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, last sequence update)
DT 10-OCT-2003 (Rel. 42, last annotation update)
DE Nitrogen regulatory protein P-II.
GN GLNB.
OS Klebsiella oxyloca.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Klebsiella.
CX NCBI_TaxID=571;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M5al;
RX MEDLINE=89201233; PubMed=2907369;
RA Hoitel A., Merrick M.; Klebsiella pneumoniae glnB gene: nucleotide
RT "Identification of the Klebsiella pneumoniae glnB gene: nucleotide
RL sequence of wild-type and mutant alleles."
MO. Gen. Genet. 215:134-138(1988).
CC -I- FUNCTION: IN NITROGEN-LIMITING CONDITIONS, WHEN THE RATIO OF GLN
TO 2-KETOGLUTARATE DECREASES, P-II IS URIDYLATED TO P-II-UMP.
P-II-UMP ALLOWS THE DEADENYLATION OF GLUTAMINE SYNTHETASE (GS),
THUS ACTIVATING THE ENZYME. CONVERSELY, IN NITROGEN EXCESS P-II
IS URIDYLATED AND PROMOTES THE ADENYLATION OF GS.
-I- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE GS
GENE (GLNA). P-II PREVENTS NR-II CATALYZED CONVERSION OF NR-I TO
NR-I-PHOSPHATE, THE TRANSCRIPTIONAL ACTIVATOR OF GLNA. WHEN P-II
IS URIDYLATED TO P-II-UMP, THESE EVENTS ARE REVERSED.
-I- SUBUNIT: Homotrimer (By similarity).
-I- SIMILARITY: Belongs to the P(II) protein family.

TO 2-KETOGLUTARATE DECREASES, P-II IS URIDYLATED TO P-II-UMP.
P-II-UMP ALLOWS THE DEADENYLATION OF GLUTAMINE SYNTHETASE (GS),
THUS ACTIVATING THE ENZYME. CONVERSELY, IN NITROGEN EXCESS P-II
IS URIDYLATED AND PROMOTES THE ADENYLATION OF GS.
-I- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE GS
GENE (GLNA). P-II PREVENTS NR-II CATALYZED CONVERSION OF NR-I TO
NR-I-PHOSPHATE, THE TRANSCRIPTIONAL ACTIVATOR OF GLNA. WHEN P-II
IS URIDYLATED TO P-II-UMP, THESE EVENTS ARE REVERSED.
-I- SUBUNIT: Homotrimer (By similarity).
-I- SIMILARITY: Belongs to the P(II) protein family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL: X51499; CAA35867.1; -
EMBL: M26107; -; NOT_ANNOTATED_CDS.
PIR: S13078; S13078.
HSSP: P05826; 2PII.
InterPro: IPR002187; PII_glnB.
InterPro: IPR002332; PII_glnB_UMP_S.
Pfam: PF00543; P-II; 1.
PRINTS: PR00340; PIIGLNB.
ProDom: PD001194; PII_glnB; 1.
PROSITE: PS00496; PII_GLNBL_UMP; 1.
PROSITE: PS00638; PII_GLNBL_CTR; 1.
KW Transcription regulation; Nitrogen fixation.
FT BINDING 51 UMP.
SQ SEQUENCE 112 AA; 12371 MW; B6258EC9C4B63871 CRC64;

Query Match 50.5%; Score 281.5; DB 1; Length 112;
Best Local Similarity 49.1%; Pred. No. 1.5e-19;
Matches 54; Conservative 26; Mismatches 29; Indels 1; Gaps 1;

QY 3 KVEATLRPRVSVSALLKIGIRGVTVSDVRGFGAQQGSTRFGQGSFSEDFKFAKVKM 62
DB 3 KIEAIRPKLDEVEALHEVGIKGVTVGKGRQKGTLYRGAEVVD-FLPKVKI 61

QY 63 EIVSKDQVEDVIEKIEEARTGEIGDKIFLLPVSVDVIRVTRGEGDKA 112
DB 62 EVVMDSDS-VERAIEAQAAHTGRIGDKGKIFVTPVEVVRITGEGGDA 111

RESULT 8
GLNB_PORPU STANDARD; PRT; 112 AA.
ID GLNB_PORPU STANDARD; PRT; 112 AA.
AC P51254;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, last sequence update)
DT 01-OCT-1996 (Rel. 34, last annotation update)
DE Nitrogen regulatory protein P-II.
GN GLNB.
OS Porphyra purpurea.
OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
CX NCBI_TaxID=2787;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Avonport;
RX Reith M.E., Munnoland J.;
RA "Complete nucleotide sequence of the Porphyra purpurea chloroplast
RT genome."
RL Plant Mol. Biol. Rep. 13:333-335(1995).
CC -I- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE
GLUTAMINE SYNTHETASE GENE (GLNA). P-II PREVENTS NR-II CATALYZED
CONVERSION OF NR-I TO NR-I-PHOSPHATE, THE TRANSCRIPTIONAL
ACTIVATOR OF GLNA. WHEN P-II IS URIDYLATED TO P-II-UMP, THESE
EVENTS ARE REVERSED. WHEN THE RATIO OF GLN TO 2-KETOGLUTARATE
DECREASES, P-II IS URIDYLATED TO P-II-UMP, WHICH CAUSES THE

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X14012; CAA32177.1; --
CC FIR; S04377; S04377.
CC HSP; P05826; 2FII.
CC InterPro; IPR002187; PII_glnB.
CC InterPro; IPR002332; PII_glnB_UMP_S.
CC Pfam; PFC0543; P-II; 1.
CC PRINTS; PR00340; PIIglnB.
CC ProDom; PD001194; PII_glnB_1.
CC PROSITE; PS00496; PII_GlnB_UMP; 1.
CC PROSITE; PS00638; PII_GlnB_CTER; 1.
CC Transcription regulation; Nitrogen fixation.
CC BINDING 51 51 UMP.
CC MUTAGEN 50 50 E->Y: GLUTAMINE AUXOTROPHY.
CC SEQUENCE 112 AA; 12429 MW; CC15D58A2F225507 CRC64;
CC -----

Query Match 49.6%; Score 276.5; DB 1; Length 112;
Best Local Similarity 50.0%; Pred. No. 4.4e-19;
Matches 55; Conservative 26; Mismatches 28; Indels 1; Gaps 1;

Qy 3 KVEAILPWRVSVSSALLKIGVTVSDVGRFGAGGSGTERQGGSEFSDKFAVKYM 62
Db 3 KIDALLKPKLDDVREALAEEVIGTMTVEYKGRQKHLYRGAEVWD-FLPKVKI 61
Qy 63 EIVVSKQDVEDVHIIIEARTGIGDKIFLLPVSVDVIRVTRGDKA 112
Db 62 EIVVTDDVDVCTDIITATGKIGDKIFVFDVARVIRITGEEDAA 111

RESULT 10
GlnB_ECOLI
ID -GlnB_ECOLI STANDARD; PRT; 112 AA.
AC P05826;
DT 01-NOV-1998 (Rel. 09, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Nitrogen regulatory protein P-II 1.
GN GlnB OR B2553 OR C3076 OR Z3829 OR ECS3419 OR STM2561 OR STY2808 OR
GN T0295 OR SF2600 OR S2772.
OS Escherichia coli.
OS Escherichia coli O6.
OS Escherichia coli O157:H7,
OS Salmonella typhimurium,
OS Salmonella typhi, and
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 217992, 83334, 602, 601, 623;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli;
RX MEDLINE=87250488; PubMed=2885322;
RA Son H.S., Rhee S.G.;
RT "Cascade control of Escherichia coli glutamine synthetase.
RT Purification and properties of PII protein and nucleotide sequence of
RT its structural gene";
RL J. Biol. Chem. 262:8690-8695(1987).
RN [2]
RN REVISIONS TO C-TERMINUS.
RC SPECIES=E.coli;
RX MEDLINE=89201233; PubMed=2907369;
RA Holtel A., Merrick M.;
RT "Identification of the Klebsiella pneumoniae glnB gene: nucleotide
RT sequence of wild-type and mutant alleles";
RL Mol. Gen. Genet. 215:134-138(1988).
RN [3]

RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=91238719; PubMed=2034230;
RA Vasudevan S.G., Amarego W.L.F., Shaw D.C., Lilley P.E., Dixon N.E.,
RA Poole R.K.;
RT "Isolation and nucleotide sequence of the hmp gene that encodes a
RT haemoglobin-like protein in Escherichia coli K-12";
RL Mol. Gen. Genet. 226:49-58(1991).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli;
RX MEDLINE=94042920; PubMed=8226691;
RA Liu J., Magasanik B.;
RT "The glnB region of the Escherichia coli chromosome";
RL J. Bacteriol. 175:7441-7449(1993).
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner P.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12";
RL Science 277:1453-1474(1997).
RN [6]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=97349980; PubMed=9205837;
RA Yanamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
RA Oshima T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasubram S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horiuchi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT - K12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features";
RL DNA Res. 4:91-113(1997).
RN [7]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O6:HI / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.F., Domeneberg M.S., Blattner P.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
RN [8]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=221074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Aantatheraman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner P.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RL Nature 409:529-533(2001).
RN [9]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12";
RL DNA Res. 8:11-22(2001).

RN SEQUENCE OF 1-12 FROM N.A.
 RP SPECIES=E.coli; STRAIN=K12 / W3110;
 RC MEDLINE=940218640; PubMed=8412694;
 RA van Heeswijk W.C., Rabenberg M., Westerhoff H.V., Kahn D.;
 RX "The genes of the glutamine synthetase adenylation cascade are not
 RT regulated by nitrogen in *Escherichia coli*.";
 RL Mol. Microbiol. 9:443-458(1993).
 RN SEQUENCE FROM N.A.
 RP SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RC MEDLINE=21534948; PubMed=11677609;
 RX McClelland M., Sanderson K.B., Spieth J., Clifton S.W., Letrelle P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium
 LT2.";
 RL Nature 413:852-856(2001).
 RN SEQUENCE FROM N.A.
 RP SPECIES=S.typhi; STRAIN=CT18;
 RC MEDLINE=21534947; PubMed=11677608;
 RX Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatina M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor T.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
 RA Krogh A., Larsen T.S., Leather S., Moulé S., O'Gaora P., Parry C.,
 RA Quail M.A., Rutherford K., Simmonds M., Skellern J., Stevens K.,
 RA Whitehead S., Barrrell B.G.;
 RT "Complete genome sequence of a multiple drug resistant *Salmonella*
 enterica serovar Typhi CT18.";
 RL Nature 413:848-852(2001).
 RN SEQUENCE FROM N.A.
 RP SPECIES=S.typhi; STRAIN=Ty2 / ATCC 700931;
 RC MEDLINE=22531367; PubMed=12644504;
 RX Deng W., Liu S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 RA Burland V., Kodoyanni V., Schwartz D.C., Blattner F.R.;
 RT "Comparative genomics of *Salmonella enterica* serovar Typhi strains Ty2
 and CT18.";
 RL J. Bacteriol. 185:2330-2337(2003).
 RN SEQUENCE FROM N.A.
 RP SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
 RC MEDLINE=22272406; PubMed=12384590;
 RX Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
 RA Yu J.;
 RT "Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity
 through comparison with genomes of *Escherichia coli* K12 and O157.";
 RL Nucleic Acids Res. 30:4432-4441(2002).
 RN SEQUENCE FROM N.A.
 RP SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;
 RC MEDLINE=2250274; PubMed=12704152;
 RX Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
 RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
 MAU B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
 RA Schwartz D.C., Blattner F.R.;
 RT "Complete genome sequence and comparative genomics of *Shigella*
 flexneri serotype 2a strain 2457T.";
 RL Infect. Immun. 71:2775-2786(2003).
 RN REVIEW.
 RP SPECIES=E.coli;
 RC MEDLINE=90089474; PubMed=2574599;
 RA Magasanik B.;
 RT "Regulation of transcription of the *glnALG* operon of *Escherichia coli*

RT by protein phosphorylation.";
 RL Biochimie 71:1005-1012(1989).
 RN CRYSTALLIZATION, AND SUBUNITS.
 RP SPECIES=E.coli;
 RC MEDLINE=94123764; PubMed=8293810;
 RX Vasudevan S.G., Gedy C., Dixon N.E., Cheah E., Carr P.D.,
 RA Suffolk P.M., Jeffrey P.D., Ollis D.L.;
 RT "Escherichia coli PII protein: purification, crystallization and
 oligomeric structure.";
 RL FEBS Lett. 337:255-258(1994).
 RN X-RAY CRYSTALLOGRAPHY (2.7 ÅNGSTROMS).
 RP SPECIES=E.coli;
 RC MEDLINE=95171116; PubMed=7866749;
 RX Cheah E., Carr P.D., Suffolk P.M., Vasudevan S.G., Dixon N.E.,
 RA Ollis D.L.;
 RT "Structure of the *Escherichia coli* signal transducing protein PII.";
 RL Structure 2:981-990(1994).
 Query Match 49.3%; Score 274.5; DB 1; Length 112;
 Best Local Similarity 50.0%; Pred. No. 6.8e-19; Indels 1; Gaps 1;
 Matches 55; Conservative 25; Mismatches 29;
 QY 3 KVEAILRPRVRSOVSSALLKIGIRGVTVSDVRFGAQQGSTERQGSSEFSEDFVAKVM 62
 DB 3 KIDALLKPKLLDDVRELAIEVGITGVTEVKGGRQKGTLYRGAEYMWVDFLPKVKI 61
 QY 63 EIVVSKQDQEVIEKIEBEARTGEIGDKIFLLPVSQVIRVIRGERGDKA 112
 DB 62 EIVVSDIVTCVDTIIRTAQTKIGDKIFVDFVARVIRVIRTEGDDAA 111
 RESULT 11
 GLNB RHORU STANDARD; PRT; 112 AA.
 ID GLNB RHORU STANDARD; PRT; 112 AA.
 AC Q53044;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Nitrogen regulatory protein P-II.
 GN GLNB.
 OS Rhodospirillum rubrum.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
 OC Rhodospirillaceae; Rhodospirillum.
 OC NCBI_TaxID=1085;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 11170 / S1;
 RX MEDLINE=96254013; PubMed=8704966;
 RA Johansson M., Nordlund S.;
 RT "Transcription of the *glnB* and *glnA* genes in the photosynthetic
 bacterium *Rhodospirillum rubrum*.";
 RL Microbiology 142:1265-1272(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 11170 / S1;
 RA Zhang Y., Ludden P.W., Roberts G.P.;
 RT Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE
 CC GLUTAMINE SYNTHETASE GENE (*GLNA*). P-II PREVENTS NR-II CATALYZED
 CC CONVERSION OF NR-I TO NR-I-PHOSPHATE, THE TRANSCRIPTIONAL
 CC ACTIVATOR OF *GLN-I* IS URIDYLATED TO P-II-UMP, THESE
 CC EVENTS ARE REVERSED. WHEN THE RATIO OF *GLN-I* TO 2-KETOGLUTARATE
 CC DECREASES, P-II IS URIDYLATED TO P-II-UMP, WHICH CAUSES THE
 CC DEADENYLATION OF GLUTAMINE SYNTHETASE, SO ACTIVATING THE ENZYME.
 CC -1- SUBUNIT: Homotrimer (By similarity).
 CC -1- SIMILARITY: Belongs to the P(II) protein family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: U32718; AAC21999.1; --
PIR: F64062; F64062.
HSP: P05826; 2P11.
TIGR: H10337; --
InterPro: IPR002187; PII_glnB.
InterPro: IPR002332; PII_glnB_UMP_S.
Pfam: PF00543; P-II; 1.
PRINTS: PR00340; PIIGLNB.
ProDom: PD001194; PII_glnB; 1.
PROSITE: PS00496; PII_glnB_UMP; 1.
PROSITE: PS00638; PII_glnB_CTR; 1.
Transcription regulation; Nitrogen fixation; Complete proteome.
KW BINDING 51 51 UMP (BY SIMILARITY).
FT BINDING 51 51 UMP (BY SIMILARITY).
SQ SEQUENCE 112 AA; 12640 MW; B8CD67EFA9381D61 CRC64;

EMBL: X71659; CAA50650.1; --
PIR: S33180; S33180.
HSP: P38504; 1GNK.
InterPro: IPR002187; PII_glnB.
InterPro: IPR002332; PII_glnB_UMP_S.
Pfam: PF00543; P-II; 1.
PRINTS: PR00340; PIIGLNB.
ProDom: PD001194; PII_glnB; 1.
PROSITE: PS00496; PII_glnB_UMP; 1.
PROSITE: PS00638; PII_glnB_CTR; 1.
Transcription regulation; Nitrogen fixation.
KW BINDING 51 51 UMP (BY SIMILARITY).
FT BINDING 51 51 UMP (BY SIMILARITY).
SQ SEQUENCE 112 AA; 12100 MW; B278486AC9EB17D3 CRC64;

Query Match 48.2%; Score 269.5; DB 1; Length 112;
Best Local Similarity 46.4%; Pred. No. 2e-18;
Matches 51; Conservative 28; Mismatches 30; Indels 1; Gaps 1;
3 KVEAILRPRVSVSALLKIGIRGVTVDVDFGAGGSGTERQGGSEDFKFAVKVM 62
3 KIEAIKPKLDEKVAEVAIVSASRTDKIDGKIFIPVEQAIRINTGETEDA 111
63 EIVVSKDQVEDVIEKIIEARTGEIGDGKIFLLPVSVDVIRVTGERGDKA 112
62 EVVLPDELVDQCEIALEITATQTKIGDKIFVYHVERAIRIRTEENEDA 111

Query Match 48.4%; Score 269.5; DB 1; Length 112;
Best Local Similarity 46.4%; Pred. No. 2e-18;
Matches 51; Conservative 28; Mismatches 30; Indels 1; Gaps 1;
3 KVEAILRPRVSVSALLKIGIRGVTVDVDFGAGGSGTERQGGSEDFKFAVKVM 62
3 KIEAIKPKLDEKVAEVAIVSASRTDKIDGKIFIPVEQAIRINTGETEDA 111
63 EIVVSKDQVEDVIEKIIEARTGEIGDGKIFLLPVSVDVIRVTGERGDKA 112
62 EVVLPDELVDQCEIALEITATQTKIGDKIFVYHVERAIRIRTEENEDA 111

RESULT 15
GLNB_RHILO
ID GLNB_RHILO STANDARD; PRT; 112 AA.
AC Q98N18;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Nitrogen regulatory protein P-II.
GN GLNB OR ML10345
OS Rhizobium loti (Mesorhizobium loti)
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFP303099;
MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,
Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
Takeuchi C., Yamada M., Tabata S.;
RA "Complete genome structure of the nitrogen-fixing symbiotic bacterium
Mesorhizobium loti.";
RT DNA Res. 7:331-338(2000).
RL
CC -1- FUNCTION: IN NITROGEN-LIMITING CONDITIONS, WHEN THE RATIO OF GLN
CC TO 2-KETOGLUTARATE DECREASES, P-II IS URIDYLATED TO P-II-UMP. P-
CC II-UMP ALLOWS THE DEADENYLATION OF GLUTAMINE SYNTHETASE (GS),
CC THUS ACTIVATING THE ENZYME. CONVERSELY, IN NITROGEN EXCESS P-II
CC IS DEURIDYLATED AND PROMOTES THE ADENYLATION OF GS (BY
CC SIMILARITY).
CC -1- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE GS
CC GENE (GLNA). P-II PREVENTS NR-II CATALYZED CONVERSION OF NR-I TO
CC NR-I-PHOSPHATE, THE TRANSCRIPTIONAL ACTIVATOR OF GLNA. WHEN P-II
CC IS URIDYLATED TO P-II-UMP, THESE EVENTS ARE REVERSED (BY
CC SIMILARITY).
CC -1- SUBUNIT: Homotrimer (By similarity).

RESULT 14
GLNB_HAEIN
ID GLNB_HAEIN STANDARD; PRT; 112 AA.
AC P43795;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Nitrogen regulatory protein P-II.
GN GLNB OR H10337.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
Weidman J.P., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
Venter J.C.;
RA "Whole-genome random sequencing and assembly of Haemophilus influenzae
Rd.";
RT
RL Science 269:496-512(1995).
CC -1- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE
CC GLUTAMINE SYNTHETASE GENE (GLNA). P-II PREVENTS NR-II CATALYZED
CC CONVERSION OF NR-I TO NR-I-PHOSPHATE, THE TRANSCRIPTIONAL
CC ACTIVATOR OF GLNA. WHEN P-II IS URIDYLATED TO P-II-UMP, THESE
CC EVENTS ARE REVERSED. WHEN THE RATIO OF GLN TO 2-KETOGLUTARATE
CC DECREASES, P-II IS URIDYLATED TO P-II-UMP, WHICH CAUSES THE
CC DEADENYLATION OF GLUTAMINE SYNTHETASE BY GLNE, SO ACTIVATING THE
CC ENZYME (BY SIMILARITY).
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- PTM: Uridylated/deuridylated by glnd (By similarity).
CC -1- SIMILARITY: Belongs to the P(II) protein family.

Fri May 28 14:57:10 2004

us-09-756-541-2.rsp

```

CC -!- SIMILARITY: Belongs to the P(II) protein family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF002994; BAB47945.1; -.
CC InterPro: IPR002187; PII_glnB.
CC InterPro: IPR002332; PII_GlnB_UMP_S.
CC Pfam: PF00543; P-II: 1.
CC PRINTS; PR00340; PII_GlnB.
CC PRODOM; PD001194; PII_GlnB; 1.
CC PROSITE; PS00496; PII_GlnB_UMP; 1.
CC PROSITE; PS00638; PII_GlnB_CTER; 1.
CC Transcription regulation: Nitrogen fixation; Complete proteome.
KW BINDING 51 UMP (BY SIMILARITY).
FT BINDING 51 UMP (BY SIMILARITY).
SQ SEQUENCE 112 AA; 12200 MW; CF0P60FE7185465A CRC64;

Query Match 48.0%; Score 267.5; DB 1; Length 112;
Best Local Similarity 47.7%; Pred. No. 3.1e-18;
Matches 51; Conservative 27; Mismatches 28; Indels 1; Gaps 1;

QY 3 KVEALLRPWRVSOVSSALLKIGIRGVTVSDVRGFGAQQSGSTERQGSSEFSEDFKFAKVKM 62
Db 3 KIEALLKPFKLDDEVKEALQEAQLQSIIVTEAKGFGKQKHTLYRGAEYVVD-FLPKVKI 61

QY 63 EIVVSKDQVEDVIRKILIEARTGIGDGKIFLLPVSDVIRVTGERG 109
Db 62 EVVLGDADAVEGAIEAIRKAAQTGRIGDGKIEFVSNIEVVRITGETG 108

Search completed: May 24, 2004, 19:13:35
Job time : 11.5 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 24, 2004, 19:09:38 ; Search time 36 Seconds
(without alignments)
990.377 Million cell updates/sec

Title: US-09-756-541-2
Perfect score: 557
Sequence: 1 FYKVEAILRPRWVSQVSSAL.....LLPVSIVIRVETGERGDAE 113

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	557	100.0	197	10 Q9ZST5	Q9ZST5 ricinus com
2	509	91.4	196	10 Q9ZST4	Q9ZST4 arabidopsis
3	492	88.3	194	10 Q9A14	Q9A14 medicago sa
4	315.5	56.6	112	16 Q9L422	Q9L422 anabaena sp
5	313.5	56.3	112	2 Q9LAQ3	Q9LAQ3 synechococc
6	311.5	55.9	112	16 Q8DLA5	Q8DLA5 synechococc
7	303.5	54.5	112	16 Q9JWC4	Q9JWC4 neisseria m
8	302.5	54.3	112	16 Q9JXK6	Q9JXK6 neisseria m
9	292.5	54.3	112	16 Q7VSR4	Q7VSR4 prochloroco
10	285.5	53.1	112	2 Q8GQS4	Q8GQS4 acetobacter
11	285.5	53.1	112	16 Q7U8Z7	Q7U8Z7 synechococc
12	282.5	52.5	112	16 Q7VA51	Q7VA51 prochloroco
13	287.5	51.6	108	2 Q8KUJ2	Q8KUJ2 neisseria m
14	286.5	51.4	112	2 Q9EZQ2	Q9EZQ2 azarcus sp
15	286.5	51.4	112	2 Q9L400	Q9L400 prochloroco
16	286.5	51.4	112	16 Q7V025	Q7V025 prochloroco

17	285.5	51.3	112	16 Q7VPK6	Q7VPK6 haemophilus
18	283.5	50.9	114	16 Q87MF6	Q87MF6 vibrio para
19	282.5	50.7	114	16 Q8DBE2	Q8DBE2 vibrio vuln
20	282.5	50.7	133	16 Q7UJ21	Q7UJ21 rhodospirill
21	281.5	50.5	105	2 Q8KUJ5	Q8KUJ5 neisseria m
22	279.5	50.2	112	16 Q8FP24	Q8FP24 corynebacte
23	278.5	50.2	112	16 Q8EBJ6	Q8EBJ6 shewanella
24	278.5	50.0	112	2 Q31188	Q31188 rhodobacter
25	278.5	50.0	112	16 Q8ZCQ9	Q8ZCQ9 yersinia pe
26	277.5	49.8	114	16 Q9KPK3	Q9KPK3 vibrio chol
27	276.5	49.6	112	2 Q9AMW9	Q9AMW9 azospirillu
28	275.5	49.5	112	16 Q7WJ14	Q7WJ14 bordetella
29	275.5	49.5	112	16 Q7WAE1	Q7WAE1 bordetella
30	275.5	49.5	112	16 Q7VVG9	Q7VVG9 bordetella
31	272.5	48.9	112	2 P70731	P70731 azospirillu
32	272.5	48.9	112	2 P94125	P94125 azorhizobiu
33	271.5	48.7	101	2 Q8KHS5	Q8KHS5 neisseria m
34	270.5	48.6	101	2 Q8KUJ0	Q8KUJ0 neisseria m
35	269.5	48.4	112	16 Q9X705	Q9X705 corynebacte
36	269.5	48.4	100	2 Q8KI77	Q8KI77 neisseria m
37	269.5	48.4	100	2 Q8KUJ4	Q8KUJ4 neisseria m
38	268.5	48.2	98	2 Q8KUJ3	Q8KUJ3 neisseria m
39	268.5	48.2	99	2 Q8KI58	Q8KI58 neisseria m
40	268.5	48.2	100	2 Q8KH45	Q8KH45 neisseria m
41	268.5	48.2	100	2 Q8KH45	Q8KH45 neisseria m
42	268.5	48.2	157	16 Q8UEI8	Q8UEI8 agrobacteri
43	267.5	48.0	98	2 Q8KUJ8	Q8KUJ8 neisseria m
44	267.5	48.0	99	2 Q8KH89	Q8KH89 neisseria m
45	267.5	48.0	112	2 P94852	P94852 herbaspiril

ALIGNMENTS

RESULT 1

Q9ZST5 PRELIMINARY: PRT; 197 AA.
AC Q9ZST5;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE PII protein (Fragment).
OS Ricinus communis (Castor bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Malpighiales; Euphorbiaceae; Acalyphoideae; Acalyphaceae;
OC Ricinus.
OX NCBI_TaxID=3988;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99030678; PubMed=9811909;
RA Hsieh M.H., Lam H.M., van de Loo F.J., Coruzzi G.;
RT "A PII-like protein in Arabidopsis: putative role in nitrogen
sensing";
RL Proc. Natl. Acad. Sci. U.S.A. 95:13965-13970 (1998).
DR EMBL; AF095454; AAC78332.1; -;
DR HSSP; P05826; 2PII.
DR GO; GO:0030234; P:enzyme regulator activity; IEA.
DR GO; GO:0006808; P:regulation of nitrogen utilization; IEA.
DR InterPro; IPR002187; PII_glnB.
DR Pfam; PF00543; P-II; 1.
DR PRINTS; PR00340; PIIGLNB.
DR ProDom; PD001194; PII_glnB; 1.
DR PROSITE; PS00638; PII_GLNB_CTER; 1.
FT NCN TER
SQ SEQUENCE 197 AA; 21781 MW; 2D18A50F15406B20 CRC64;

Query Match 100.0%; Score 557; DB 10; Length 197;
Best Local Similarity 100.0%; Pred. No. 2.3e-45;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FYKVEAILRPRWVSQVSSALLKIGIRGVTVSDVRFGAGGSTERGGSEFEDFKVAKV 60
|||||

Db 72 FYKVEAILRPMRVSVSSALLKIGIRGVTVDVRGFGAGGSTERQGGSEFSEDKFVAKV 131
QY 61 KMEIVVSKDOVEDVIEKTIERTARTGEIGDKIFLLPVSVDVIRVTRGEGDKAE 113
Db 132 KMEIVVSKDOVEDVIEKTIERTARTGEIGDKIFLLPVSVDVIRVTRGEGDKAE 184
RESULT 2
Q9ZST4 PRELIMINARY; PRT; 196 AA.
AC Q9ZST4;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE PII protein (P II nitrogen sensing protein GLB I) (At4g01900).
GN T7B11.16 OR AT4G01900
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=99030678; PubMed=9811909;
RA Hsieh M.H., Lam H.M., Van de Loo P.J., Coruzzi G.;
RT "A PII-like protein in Arabidopsis: putative role in nitrogen
sensing.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:13965-13970(1998).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Huang E.N., Nascimento L., de la Bastide M., Habermann K., Vil M.D.,
RA Preston R.R., Spiegel L.A., See L.H., Shah R., Matero A.,
RA O'Shaughnessy A., Rodriguez M., Shekher M., Swaby I., Schutz K.,
RA Parnell L.D., Dedhia N.N., McCombie W.R.;
RA "Arabidopsis thaliana BAC T7B11 from chromosome IV near 10 cm.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
SEQUENCE FROM N.A.
RA Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,
RA Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,
RA Shekher M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,
RA Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Shinn P., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Hsuan V.W.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G., Yuan S.,
RA Shinzaki K., Davis R.W., Theologis A., Ecker J.R.;
RL "Arabidopsis ORF clones.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF095455; AAC78333.1; -;
DR EMBL; AC007138; AAD2652.1; -;
DR EMBL; AL161493; CAB80683.1; -;
DR EMBL; BT005209; AAO63273.1; -;
DR PIR; D85024; D85024.
DR HSSP; P05826; 2PII.
DR GO; GO:0030234; P:enzyme regulator activity; IEA.
DR GO; GO:0006808; P:regulation of nitrogen utilization; IEA.
DR InterPro; IPR002187; PII_glnB.
DR Pfam; PF00543; P-II; 1.
DR PRINTS; PD00340; PIIGLNB.
DR ProDom; PD001194; PII_glnB.
DR PROSITE; PS00638; PII_GlnB_CTER; 1.
SEQUENCE 196 AA; 21275 MW; FE740EA66776F157 CRC64;

Query Match 91.4%; Score 509; DB 10; Length 196;
Best Local Similarity 90.3%; Pred. No. 8.9e-41;
Matches 102; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
QY 1 FYKVEAILRPMRVSVSSALLKIGIRGVTVDVRGFGAGGSTERQGGSEFSEDKFVAKV 60
Db 74 FYKVEAILRPMRVSVSSALLKIGIRGVTVDVRGFGAGGSTERQGGSEFSEDKFVAKV 133
QY 61 KMEIVVSKDOVEDVIEKTIERTARTGEIGDKIFLLPVSVDVIRVTRGEGDKAE 113
Db 134 KMEIVVSKDOVEDVIEKTIERTARTGEIGDKIFLLPVSVDVIRVTRGEGDKAE 186
RESULT 3
Q9ARI4 PRELIMINARY; PRT; 194 AA.
AC Q9ARI4;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE PII protein.
GN GLNB.
OS Medicago sativa (Alfalfa).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
OX NCBI_TaxID=3879;
RN [1]
SEQUENCE FROM N.A.
RA Garcia-Ibáñez D., Sengupta-Gopalan C.;
RL "Characterization of PII (GLNB) in alfalfa.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY027892; AAK16221.1; -;
DR HSSP; P38504; 1GNK.
DR GO; GO:0030234; P:enzyme regulator activity; IEA.
DR GO; GO:0006808; P:regulation of nitrogen utilization; IEA.
DR InterPro; IPR002187; PII_glnB.
DR Pfam; PF00543; P-II; 1.
DR PRINTS; PR00340; PIIGLNB.
DR ProDom; PD001194; PII_glnB.
DR PROSITE; PS00638; PII_GlnB_CTER; 1.
SEQUENCE 194 AA; 21413 MW; 23FA623FF4D97450 CRC64;
Query Match 88.3%; Score 492; DB 10; Length 194;
Best Local Similarity 85.8%; Pred. No. 3.7e-39;
Matches 97; Conservative 8; Mismatches 8; Indels 0; Gaps 0;
QY 1 FYKVEAILRPMRVSVSSALLKIGIRGVTVDVRGFGAGGSTERQGGSEFSEDKFVAKV 60
Db 70 FYKVEAILRPMRVSVSSALLKIGIRGVTVDVRGFGAGGSTERQGGSEFSEDKFVAKV 129
QY 61 KMEIVVSKDOVEDVIEKTIERTARTGEIGDKIFLLPVSVDVIRVTRGEGDKAE 113
Db 130 KMEIVVSKDOVEDVIEKTIERTARTGEIGDKIFLLPVSVDVIRVTRGEGDKAE 182
RESULT 4
Q9L422 PRELIMINARY; PRT; 112 AA.
AC Q9L422;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE PII protein (Nitrogen regulatory protein P-II).
GN GLNB OR ALL2319.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=PCC 7120;
RA Gonzalez L., Phalip V., Zhang C.C.;

Neisseria meningitidis (serogroup A).
Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
C
Neisseriaceae; Neisseria.
X NCBI_TaxID=65699;
[!]
SEQUENCE FROM N.A.
P STRAIN=22491 / Serogroup A / Serotype 4A;
C MEDLINE=20222556; PubMed=10761919;
C PARKILL J., Achtman M., James K.D., Bentley S.D., Churcher C.,
A Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
Davies R.M., Davis P., Devlin K., Fellwell T., Hamlin N., Holroyd S.,
Jagels K., Leachner S., Moule S., Mangall K., Quail M.A.,
Rajadream M.A., Rutherford K.M., Simmonds M., Skelton J.,
Whitehead S., Spratt B.G., Barrall B.G.;
A "Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis Z2491";
T Nature 404:592-506(2000).
L EMBL; AL162753; CAB83744.1; -.
R HSP; P05826; 2PII.
R PSR; H81961; H81961.
R GO; GO:0006808; P:enzyme regulator activity; IEA.
R GO; GO:0006808; P:regulation of nitrogen utilization; IEA.
R InterPro; IPR002187; PII_glnB_UMP_S.
R Pfam; PF00543; P-II; 1.
R PRINTS; PR00340; PIIGLN.B.
R ProDom; PD001194; PII_glnB; 1.
R PROSITE; PS00638; PII_GLN_B_CTER; 1.
R PROSITE; PS00496; PII_GLN_UMP; 1.
R Complete proteome.
K Complete proteome.
W Complete proteome.
SQ SEQUENCE 112 AA; 12281 MW; CD893CF24A1A485F CRC64;

Query Match 54.5%; Score 303.5; DB 16; Length 112;
Best Local Similarity 54.5%; Pred. No. 2e-21;
Matches 60; Conservative 25; Mismatches 24; Indels 1; Gaps 1;
ID QTVSR4 PRELIMINARY; PRT; 112 AA.

AC QTVSR4
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nitrogen regulatory protein P-II.
GN GLNB OR PWT1481.
OS Escherichia coli (strain ATCC 8739) [NCBI TaxID=562].
OC Bacteria; Gammaproteobacteria; Enterobacteriaceae;
OC Escherichia coli.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20222556; PubMed=10761919;
RA Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,
RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
RA Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,
RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
RA Webb E.A., Zinger E.R., Chisholm S.W.;
RT "Genomic divergence in two Escherichia coli ecotypes reflects oceanic
RT niche differentiation.";
RL Nature 424:1042-1047(2003).
DR EMBL; BX572099; CAE21656.1; -.
KW Complete proteome.
SQ SEQUENCE 112 AA; 12281 MW; CD893CF24A1A485F CRC64;

Query Match 54.5%; Score 302.5; DB 16; Length 112;
Best Local Similarity 53.6%; Pred. No. 2.5e-21;
Matches 59; Conservative 26; Mismatches 24; Indels 1; Gaps 1;
ID QTVSR4 PRELIMINARY; PRT; 112 AA.

AC QTVSR4
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nitrogen regulatory protein P-II.
GN GLNB OR PWT1481.
OS Escherichia coli (strain ATCC 8739) [NCBI TaxID=562].
OC Bacteria; Gammaproteobacteria; Enterobacteriaceae;
OC Escherichia coli.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20222556; PubMed=10761919;
RA Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,
RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
RA Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,
RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
RA Webb E.A., Zinger E.R., Chisholm S.W.;
RT "Genomic divergence in two Escherichia coli ecotypes reflects oceanic
RT niche differentiation.";
RL Nature 424:1042-1047(2003).
DR EMBL; BX572099; CAE21656.1; -.
KW Complete proteome.
SQ SEQUENCE 112 AA; 12281 MW; CD893CF24A1A485F CRC64;

Query Match 54.3%; Score 302.5; DB 16; Length 112;
Best Local Similarity 53.6%; Pred. No. 2.5e-21;
Matches 59; Conservative 26; Mismatches 24; Indels 1; Gaps 1;
ID QTVSR4 PRELIMINARY; PRT; 112 AA.

AC QTVSR4
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nitrogen regulatory protein P-II.
GN GLNB OR PWT1481.
OS Escherichia coli (strain ATCC 8739) [NCBI TaxID=562].
OC Bacteria; Gammaproteobacteria; Enterobacteriaceae;
OC Escherichia coli.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20222556; PubMed=10761919;
RA Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,
RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
RA Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,
RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
RA Webb E.A., Zinger E.R., Chisholm S.W.;
RT "Genomic divergence in two Escherichia coli ecotypes reflects oceanic
RT niche differentiation.";
RL Nature 424:1042-1047(2003).
DR EMBL; BX572099; CAE21656.1; -.
KW Complete proteome.
SQ SEQUENCE 112 AA; 12281 MW; CD893CF24A1A485F CRC64;

Query Match 54.3%; Score 302.5; DB 16; Length 112;
Best Local Similarity 53.6%; Pred. No. 2.5e-21;
Matches 59; Conservative 26; Mismatches 24; Indels 1; Gaps 1;
ID QTVSR4 PRELIMINARY; PRT; 112 AA.

AC QTVSR4
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nitrogen regulatory protein P-II.
GN GLNB OR PWT1481.
OS Escherichia coli (strain ATCC 8739) [NCBI TaxID=562].
OC Bacteria; Gammaproteobacteria; Enterobacteriaceae;
OC Escherichia coli.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20222556; PubMed=10761919;
RA Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,
RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
RA Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,
RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
RA Webb E.A., Zinger E.R., Chisholm S.W.;
RT "Genomic divergence in two Escherichia coli ecotypes reflects oceanic
RT niche differentiation.";
RL Nature 424:1042-1047(2003).
DR EMBL; BX572099; CAE21656.1; -.
KW Complete proteome.
SQ SEQUENCE 112 AA; 12281 MW; CD893CF24A1A485F CRC64;

Query Match 54.3%; Score 302.5; DB 16; Length 112;
Best Local Similarity 53.6%; Pred. No. 2.5e-21;
Matches 59; Conservative 26; Mismatches 24; Indels 1; Gaps 1;
ID QTVSR4 PRELIMINARY; PRT; 112 AA.

AC QTVSR4
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nitrogen regulatory protein P-II.
GN GLNB OR PWT1481.
OS Escherichia coli (strain ATCC 8739) [NCBI TaxID=562].
OC Bacteria; Gammaproteobacteria; Enterobacteriaceae;
OC Escherichia coli.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20222556; PubMed=10761919;
RA Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,
RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
RA Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,
RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
RA Webb E.A., Zinger E.R., Chisholm S.W.;
RT "Genomic divergence in two Escherichia coli ecotypes reflects oceanic
RT niche differentiation.";
RL Nature 424:1042-1047(2003).
DR EMBL; BX572099; CAE21656.1; -.
KW Complete proteome.
SQ SEQUENCE 112 AA; 12281 MW; CD893CF24A1A485F CRC64;

Query Match 54.3%; Score 302.5; DB 16; Length 112;
Best Local Similarity 53.6%; Pred. No. 2.5e-21;
Matches 59; Conservative 26; Mismatches 24; Indels 1; Gaps 1;
ID QTVSR4 PRELIMINARY; PRT; 112 AA.

AC QTVSR4
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nitrogen regulatory protein P-II.
GN GLNB OR PWT1481.
OS Escherichia coli (strain ATCC 8739) [NCBI TaxID=562].
OC Bacteria; Gammaproteobacteria; Enterobacteriaceae;
OC Escherichia coli.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20222556; PubMed=10761919;
RA Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,
RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
RA Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,
RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
RA Webb E.A., Zinger E.R., Chisholm S.W.;
RT "Genomic divergence in two Escherichia coli ecotypes reflects oceanic
RT niche differentiation.";
RL Nature 424:1042-1047(2003).
DR EMBL; BX572099; CAE21656.1; -.
KW Complete proteome.
SQ SEQUENCE 112 AA; 12281 MW; CD893CF24A1A485F CRC64;

Query Match 54.3%; Score 302.5; DB

[illegible]

Fri May 28 14:57:10 2004

DE PRINTS; PR00340; PII_GlnB.
GN ProDom; PD001194; PII_glnB; 1.
OS PROSITE; PS00638; PII_GlnB_CTER; 1.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococaceae;
OC Prochlorococcus.
OK NCBI_TaxID=100363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=strain PCC 9511;
RA Palinska K.A., Loiseux-de Goer S., Blondel A., Castets A.M.,
RA Kippka R., Tandeau de Marsac N.;
RT "In spite of synthesizing a cyanobacterial-type PII protein, the
RT oceanic photosynthetic prokaryote Prochlorococcus marinus, strain PCC
RT 9511, is unable to utilize nitrate."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ271089; CAB87556.1; -
DR HSSP; P38504; 1GNK.
DR GO; GO:0030234; F:enzyme regulator activity; IEA.
DR GO; GO:0006808; P:regulation of nitrogen utilization; IEA.
DR InterPro; IPR002187; PII_glnB.
DR Pfam; PF00543; P-II; 1.
DR PRINTS; PR00340; PII_GlnB.
DR ProDom; PD001194; PII_glnB; 1.
DR PROSITE; PS00638; PII_GlnB_CTER; 1.
SQ SEQUENCE 112 AA; 12314 MW; 42B3497E6B4472CD CRC64;
Query Match 51.4%; Score 286.5; DB 2; Length 112;
Best Local Similarity 50.9%; Pred. No. 8.5e-20;
Matches 56; Conservative 25; Mismatches 28; Indels 1; Gaps 1;
QY 3 KVEAILRPMRVSVQSSALLKIGIRGVTVSDVRGFGAGGSTERQGGSEFSEDKFVAKVM 62
DB 3 KIEAILRPFKLDVREALSEVGIAGLTVTVKVGFGKQKHTLYRGAEYVVD-FLPKIKV 61
QY 63 EIVSVKQVEDVIEKIIIEARTGEIGDKIFLLPVSDVIRVGTGERGDKA 112
DB 62 EVVVENKSVSVIDAIAEAAKTGEIGDKIFISSIDSIVRIRTGTDEEA 111
Search completed: May 24, 2004, 19:15:02
Job time : 37 secs

DR PRINTS; PR00340; PII_GlnB.
GN ProDom; PD001194; PII_glnB; 1.
OS PROSITE; PS00638; PII_GlnB_CTER; 1.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococaceae;
OC Prochlorococcus.
OK NCBI_TaxID=100363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=strain PCC 9511;
RA Palinska K.A., Loiseux-de Goer S., Blondel A., Castets A.M.,
RA Kippka R., Tandeau de Marsac N.;
RT "In spite of synthesizing a cyanobacterial-type PII protein, the
RT oceanic photosynthetic prokaryote Prochlorococcus marinus, strain PCC
RT 9511, is unable to utilize nitrate."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ271089; CAB87556.1; -
DR HSSP; P38504; 1GNK.
DR GO; GO:0030234; F:enzyme regulator activity; IEA.
DR GO; GO:0006808; P:regulation of nitrogen utilization; IEA.
DR InterPro; IPR002187; PII_glnB.
DR Pfam; PF00543; P-II; 1.
DR PRINTS; PR00340; PII_GlnB.
DR ProDom; PD001194; PII_glnB; 1.
DR PROSITE; PS00638; PII_GlnB_CTER; 1.
DR PROSITE; PS00496; PII_GlnB_UMP; 1.
SQ SEQUENCE 108 AA; 11812 MW; BE98D250C543ACBD CRC64;
Query Match 51.6%; Score 287.5; DB 2; Length 108;
Best Local Similarity 52.8%; Pred. No. 6.5e-20;
Matches 57; Conservative 25; Mismatches 25; Indels 1; Gaps 1;
QY 5 EAILRPMRVSVQSSALLKIGIRGVTVSDVRGFGAGGSTERQGGSEFSEDKFVAKVM 64
DB 1 EADVPKFLDDVREALTEIGTGMTVSEVKGFGKQKHTLYRGAEYVVD-FLPKIKV 59
QY 65 VVSKQVEDVIEKIIIEARTGEIGDKIFLLPVSDVIRVGTGERGDKA 112
DB 60 VLADDAVERAIDVIVEVARSGKIGDKIFVLPVERAIRIRTGERSDAA 107
RESULT 14
Q9EZQ2 PRELIMINARY; PRT; 112 AA.
AC Q9EZQ2; MEDLINE=20521930; PubMed=11059654;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE P2-like signal transmitter protein GlnB.
GN GlnB.
OS Azotarcus sp. BH72.
OC Bacteria; Proteobacteria; Betaproteobacteria; Rhodocyclales;
OC Rhodocyclaceae; Azotarcus.
OX NCBI_TaxID=62928;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BH72;
RX MEDLINE=20521930; PubMed=11059654;
RA Martin D.E., Hurek T., Reinhold-Hurek B.;
RT "Occurrence of three PII-like signal transmitter proteins in the
RT diazotrophic proteobacterium Azotarcus sp. BH72."
RL Mol. Microbiol. 38:276-288(2000).
DR EMBL; AF281017; AAC40188.1; -
DR HSSP; P05826; 2PII.
DR GO; GO:0030234; F:enzyme regulator activity; IEA.
DR GO; GO:0006808; P:regulation of nitrogen utilization; IEA.
DR InterPro; IPR002187; PII_glnB.
DR InterPro; IPR002332; PII_GlnB_UMP_S.
DR Pfam; PF00543; P-II; 1.
DR ProDom; PD001194; PII_glnB; 1.
DR PROSITE; PS00638; PII_GlnB_CTER; 1.
DR PROSITE; PS00496; PII_GlnB_UMP; 1.
SQ SEQUENCE 112 AA; 12273 MW; 636D8AF1909426A7 CRC64;
Query Match 51.4%; Score 286.5; DB 2; Length 112;
Best Local Similarity 48.2%; Pred. No. 8.5e-20;
Matches 53; Conservative 30; Mismatches 26; Indels 1; Gaps 1;
QY 3 KVEAILRPMRVSVQSSALLKIGIRGVTVSDVRGFGAGGSTERQGGSEFSEDKFVAKVM 62
DB 3 KIEAILRPFKLDVREALSEVGIAGLTVTVKVGFGKQKHTLYRGAEYVVD-FLPKIKV 61
QY 63 EIVSVKQVEDVIEKIIIEARTGEIGDKIFLLPVSDVIRVGTGERGDKA 112
DB 62 EVVADDVVECTMDAVIKAAOTGKIGDKIFVTPVEQVRIRTGTETNEAA 111
RESULT 15
Q9L400 PRELIMINARY; PRT; 112 AA.
AC Q9L400;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2004, 13:35:34 ; Search time 2414.89 Seconds
(without alignments)
14663.718 Million cell updates/sec

Title: US-09-756-541-13
Perfect score: 817
Sequence: 1 CTGAAGTGTGTATAAAA.....GTTCTATAATAAAAAAAA 817

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:

- 1: gb_ba.*
- 2: gb_hgt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_hgt_hum.*
- 31: em_hgt_inv.*
- 32: em_hgt_other.*
- 33: em_hgt_mus.*
- 34: em_hgt_pln.*
- 35: em_hgt_rtd.*
- 36: em_hgt_man.*
- 37: em_hgt_vrt.*
- 38: em_sy.*
- 39: em_hgt_hum.*
- 40: em_hgt_mus.*
- 41: em_hgt_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	817	100.0	817	6	AR125589	AR125589 Sequence
2	808	98.9	844	8	AF095455	AF095455 Arabidops
3	591	72.3	591	6	AX507438	AX507438 Sequence
4	591	72.3	591	8	BT005209	BT005209 Arabidops
5	588	72.0	588	6	AR125591	AR125591 Sequence
6	286.4	35.1	897	6	AR125590	AR125590 Sequence
7	285	34.9	594	6	AR125592	AR125592 Sequence
8	285	34.9	840	8	AF095454	AF095454 Ricinus c
9	253.8	31.1	796	8	AY027892	AY027892 Medicago
10	246	30.1	120185	8	AC007138	AC007138 Arabidops
11	246	30.1	198220	8	ATCHRIV5	AT161493 Arabidops
12	245	30.0	947	8	AY442185	AY442185 Lycopersi
13	209.6	25.7	1085	8	PPI489604	AT489604 Pinus pin
14	199.4	24.4	902	8	AK068407	AK068407 Oryza sat
15	199.4	24.4	902	8	AK099152	AK099152 Oryza sat
16	109.8	13.4	384	1	FDGLNBPRT	X97327 F. diplosiph
17	107.4	13.1	300143	1	AE017165	AE017165 Prochloro
18	106.8	13.1	164921	8	AF022186	AF022186 Cyanidriu
19	106	13.0	2493	1	AS0251822	AT251822 Anabaena
20	106	13.0	341880	1	AF003589	AF003589 Nostoc sp
21	102	12.5	262202	1	EX572094	EX572094 Prochloro
22	101.6	12.4	191028	8	PPU38804	U38804 Porphyra pu
23	98.6	12.1	4348	1	AF079137	AF079137 Synechoco
24	98	12.0	686	1	AF017419	AF017419 Nostoc pu
25	97	11.9	383	1	SYOGLNB	ME2447 Synechococc
26	95.8	11.7	339	1	PMA271089	AD271089 Prochloro
27	95.8	11.7	11771	1	AE000674	AE000674 Aquifex a
28	95.2	11.7	302320	1	AP005081	AP005081 Vibrio pa
29	94.2	11.5	349746	1	EX572099	EX572099 Prochloro
30	94	11.5	299350	1	AP005370	AP005370 Thermosyn
31	93.6	11.5	302050	1	AP006568	AP006568 Gloeobact
32	93.2	11.4	300045	1	AR016803	AR016803 Vibrio vu
33	92.8	11.4	96109	6	AR408756	AR408756 Sequence
34	92.8	11.4	96109	6	AX067460	AX067460 Sequence
35	91.6	11.2	247950	1	AP005340	AP005340 Pasteurel
36	89.8	11.0	11371	1	AE006236	AE006236 Pasteurel
37	89.4	10.9	351	6	AR378407	AR378407 Sequence
38	89.4	10.9	300732	1	AR016802	AR016802 Vibrio vu
39	89	10.9	1622	1	AF120107	AF120107 Synechoco
40	88.4	10.8	452	1	SSGLNBP	X97496 Synechocyst
41	88.4	10.8	130001	1	D90915	D90915 Synechocyst
42	86.8	10.6	298950	1	AP004597	AP004597 Oceanobac
43	86.2	10.6	248650	1	AP005341	AP005341 Vibrio vu
44	85.8	10.5	375	6	AR318761	AR318761 Sequence
45	85.4	10.5	405	6	AR376903	AR376903 Sequence

ALIGNMENTS

RESULT 1
AR125589
LOCUS
DEFINITION Sequence 13 from patent US 6177275.
ACCESSION AR125589
VERSION AR125589.1 GI:14111651
KEYWORDS
SOURCE
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 817)
AUTHORS Coruzzi, G.M., Lam, H.-M. and Hsieh, M.-H.
TITLE Plant nitrogen regulatory P-III genes
JOURNAL Patent: US 6177275-A 13 23-JAN-2001;
FEATURES Location/Qualifiers

linear PAT 16-MAY-2001

RESULT 2	AF095455	AF095455	844 bp	linear	PLN 19-NOV-1998
LOCUS		Arabidopsis thaliana	P11 protein mRNA, complete cds.		
DEFINITION					
ACCESSION		AF095455			
VERSION		AF095455.1			
KEYWORDS		GI:3885942			

```

483 ATAAATTGAAGGAGCAAGGACAGGAGAGATTGGTGAATGCGCAAGATTTTGTGTTGCTGTG 542
481 ATAAATTGAAGGAGCAAGGACAGGAGAGATTGGTGAATGCGCAAGATTTTGTGTTGCTGTG 540
543 TCAGATGTCATAAGAGTTAGGACAGGTCAGCGTGGGAGAGAAAGCAGAGAGATGACTGGT 602
541 TCAGATGTCATAAGAGTTAGGACAGGTCAGCGTGGGAGAGAAAGCAGAGAGATGACTGGT 600
603 GATATGCTTTCACCGTCTTAGGAAACAAACAGAGCTCAAGATGGTTTTTTTTTTTTCAT 662
601 GATATGCTTTCACCGTCTTAGGAAACAAACAGAGCTCAAGATGGTTTTTTTTTTTTCAT 660
663 TCGGTCCTCAGATCTCGGATATATATGATGAGTCGAGTCGTTGTTGTTTTCATGTTGAA 722
661 TCGGTCCTCAGATCTCGGATATATATGATGAGTCGAGTCGTTGTTGTTTTCATGTTGAA 720
723 TCGATCAGATGTTTTTAACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 782
721 TCGATCAGATGTTTTTAACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
783 AGACATCGAATCTGTTCTTAATAAAA 810
781 AGACATCGAATCTGTTCTTAATAAAA 808

RESULT 3
AX507438
LOCUS AX507438 591 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 2133 from Patent WO216655.
ACCESSION AX507438
VERSION AX507438.1 GI:23388675
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
Harper, J.F., Kreps, J., Wang, X. and Zhu, T.
Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
Patent: WO 0216655-A 2133 28-FEB-2002;
The Scripps Research Institute (US); Syngenta Participations AG
(CH)

FEATURES
source
1. .591
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Query March 72.3%; Score 591; DB 6; Length 591;
Best Local Similarity 100.0%; Pred. No. 4.1e-130;
Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

33 ATGCGCGGTCATGACGAAACCCATCTCAATACCTCTCTCGTTTCTATTCTGATCGA 92
1 ATGCGCGGCGATGACGAAACCCATCTCAATACCTCTCTCGTTTCTATTCTGATCGA 60
93 AAGNACATTCCTCTCTGATGTCATTTGTCATTTGTCGATTCAGACATTCGCCACCA 152
61 AAGAACATTCCTCTCTGATGTCATTTGTCATTTGTCGATTCAGACATTCGCCACCA 120
153 TCTTGCCTCGATTTGGTCACAAAGTCACCGAGTAATACAGTCGTGTTTACCTGTCGT 212
121 TCTTGCCTCGATTTGGTCACAAAGTCACCGAGTAATACAGTCGTGTTTACCTGTCGT 180
213 AGTGCCCAATATCTCTGATATATTCAGACATCGAAATTTTACAGGTGAGCAATT 272
181 AGTGCCCAATATCTCTGATATATTCAGACATCGAAATTTTACAGGTGAGCAATT 240
273 GTGACAGCTGGAATCCAGCAAGTTTCATCGGCTTCTACTGAAATTCGGGATTCGAGGT 332

```

```

241 GTCGACCATCGAGAAATCCAGCAAGTTTTCATCGGCTTTTACTGAAAATCGGGATTCGAGGT 300
333 GTTACTGTTTCTGATGTGAGAGGTTTGGTGCACAGGAGGTTTCTACCGAGAGACACCGGT 392
301 GTTACTGTTTCTGATGTGAGAGGTTTGGTGCACAGGAGGTTTCTACCGAGAGACACCGGT 360
393 GGCTCTGAGTTTCTCGGAAGACAAATTTGTTGCTTAAAGTTTAAAGTGGAAATCGTTGTTAAG 452
361 GGCTCTGAGTTTCTCGGAAGACAAATTTGTTGCTTAAAGTTTAAAGTGGAAATCGTTGTTAAG 420
453 AAGACCAAGTGGAAATCTGTAATCAACACAATAATTAAGAGGAGGAGGAGAGATT 512
421 AAGACCAAGTGGAAATCTGTAATCAACACAATAATTAAGAGGAGGAGGAGAGATT 480
513 GGTGATGGAAGATTTTGTGTTGCTGTCAGATGTCATAGAGTTAGGACAGGTGAG 572
481 GGTGATGGAAGATTTTGTGTTGCTGTCAGATGTCATAGAGTTAGGACAGGTGAG 540
573 CGTGGGAGAGAGGAGAGAGATGACTGCTGATGATGCTTTCACCGTCTTAG 623
541 CGTGGGAGAGAGGAGAGAGATGACTGCTGATGATGCTTTCACCGTCTTAG 591

RESULT 4
BT005209
LOCUS BT005209 591 bp mRNA linear PLN 14-MAR-2003
DEFINITION Arabidopsis thaliana At4g01900 mRNA, complete cds.
ACCESSION BT005209
VERSION BT005209.1 GI:28950698
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 591)
Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Carninci, P.,
Chan, M.M., Chang, C.H., Dale, J.M., Hayashizaki, Y., Hsuan, V.W.,
Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,
Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,
Southwick, A., Tang, C.C., Toriumi, M., Mong, C., Wu, H.C., Yamada, K.,
Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and
Ecker, J.R.
Arabidopsis ORF clones
Unpublished
2 (bases 1 to 591)
Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Carninci, P.,
Chan, M.M., Chang, C.H., Dale, J.M., Hayashizaki, Y., Hsuan, V.W.,
Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,
Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,
Southwick, A., Tang, C.C., Toriumi, M., Mong, C., Wu, H.C., Yamada, K.,
Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and
Ecker, J.R.
Direct Submission
Submitted (14-MAR-2003) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL CDNA : 'RIKEN
Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinozaki, K.
The Salk, Stanford, PBC (SSP) Consortium members constructed and
sequenced the pN1 (ORF) clones using the RAFL cDNAs: Kim, C.J.,
Chen, H., Cheuk, R., Shinn, P., Bowser, L., Chan, M.M., Chang, C.M.,
Dale, J.M., Hsuan, V.W., Jones, T., Karlin-Neumann, G., Lam, B.,
Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J.,
Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Mong, C., Wu, H.C.,
Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A., and

```

Ecker, J.R.
Kim, C.J. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES
Source
Location/Qualifiers
1. .591
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="4"
/clone="U23463"
/ecotype="Columbia"
/note="This clone is in pUNI 51"
1. .591
/notes="p II nitrogen sensing protein GLB 1"
/codon_start=1
/product="At4g01900"
/protein_id="AA063273.1"
/db_xref="GI:28950699"
/translation="MAASMKPISITSLGFSYDRKNIAFDSCISICSGFHRSPRLD
LVTKSPNSRVLPLVWSAQISDYIPDSKFKVEIWPRIQVSSALIKIGIRGT
VSDVRGFGAQSGSTERHGSSEFSEDKFAVKVMEIVVKQDVESVINIIEGARTGEI
GDGKIPLVPSDVIRVTRGERGEKAEKWTGMLSPS"

CDS
1. .591
/notes="p II nitrogen sensing protein GLB 1"
/codon_start=1
/product="At4g01900"
/protein_id="AA063273.1"
/db_xref="GI:28950699"
/translation="MAASMKPISITSLGFSYDRKNIAFDSCISICSGFHRSPRLD
LVTKSPNSRVLPLVWSAQISDYIPDSKFKVEIWPRIQVSSALIKIGIRGT
VSDVRGFGAQSGSTERHGSSEFSEDKFAVKVMEIVVKQDVESVINIIEGARTGEI
GDGKIPLVPSDVIRVTRGERGEKAEKWTGMLSPS"

ORIGIN
Query Match 72.3%; Score 591; DB 8; Length 591;
Best Local Similarity 100.0%; Pred. No. 4.1e-130;
Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 33 ATGGCGCGTCAATGAGAAACCCATCTCAATACTTCTCGGTTCTATTCTGATCGA 92
Db 1 ATGGCGCGTCAATGAGAAACCCATCTCAATACTTCTCGGTTCTATTCTGATCGA 60
Qy 93 AAGAACATTCCTTCTCTGATTGCAATTCGATTTGTTCTGATTCAGACATTCGCGACCA 152
Db 61 AAGAACATTCCTTCTCTGATTGCAATTCGATTTGTTCTGATTCAGACATTCGCGACCA 120
Qy 153 TCTTGCCTCGATTTGGTCAAAAGTCCAGAGTAAATACAGTGTGTTTACCTGTGCTT 212
Db 121 TCTTGCCTCGATTTGGTCAAAAGTCCAGAGTAAATACAGTGTGTTTACCTGTGCTT 180
Qy 213 AGTCCCAAAATATCTTCTGATTATATCCAGACTCGAAATTTTACAAGTGGAGCAAT 272
Db 181 AGTCCCAAAATATCTTCTGATTATATCCAGACTCGAAATTTTACAAGTGGAGCAAT 240
Qy 273 GTCAGACCATGAGAAATCCAGAAAGTTCATCGGCTTTACTGAAATCGGATTCGAGT 332
Db 241 GTCAGACCATGAGAAATCCAGAAAGTTCATCGGCTTTACTGAAATCGGATTCGAGT 300
Qy 333 GTTACTGTTCTGATGTGAGAGGTTTGGTGCACAGAGGTTTCTACCGAGAGACACGGT 392
Db 301 GTTACTGTTCTGATGTGAGAGGTTTGGTGCACAGAGGTTTCTACCGAGAGACACGGT 360
Qy 393 GCCTCTGAGTTCTCGGAGACAAATTTGTTCTAAAGTTAAGATCGAAATCGTTGTTAAG 452
Db 361 GCCTCTGAGTTCTCGGAGACAAATTTGTTCTAAAGTTAAGATCGAAATCGTTGTTAAG 420
Qy 453 AAAGACCAAGTGGATCTGTAATCAACAAATAATTGAAGGAGCAAGGACGAGAGATT 512
Db 421 AAAGACCAAGTGGATCTGTAATCAACAAATAATTGAAGGAGCAAGGACGAGAGATT 480
Qy 513 GGTGATGCAAGATTTTGTGTTGCTGTGATGATGATGATGATGATGATGATGATGATGAT 572
Db 481 GGTGATGCAAGATTTTGTGTTGCTGTGATGATGATGATGATGATGATGATGATGATGAT 540
Qy 573 CGTGGGGAGAGACGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 623
Db 541 CGTGGGGAGAGAGACGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 591
RESULT 5
AR125591

LOCUS
DEFINITION
AR125591
SEQUENCE 15 from patent US 6177275.
588 bp DNA linear PAT 16-MAY-2001
AR125591.1 GI:14111653
KEYWORDS
SOURCE
ORGANISM
Unknown.
Unclassified.
REFERENCE
1 (bases 1 to 588)
Coruzzi, G.M., Lam, H.-M. and Hsieh, M.-H.
Plant nitrogen regulatory P-PII genes
Patent: US 6177275-A 15 23-JAN-2001;
Location/Qualifiers
FEATURES
1. .588
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 72.0%; Score 588; DB 6; Length 588;
Best Local Similarity 100.0%; Pred. No. 2.1e-129;
Matches 588; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 33 ATGGCGCGTCAATGAGAAACCCATCTCAATACTTCTCGGTTCTATTCTGATCGA 92
Db 1 ATGGCGCGTCAATGAGAAACCCATCTCAATACTTCTCGGTTCTATTCTGATCGA 60
Qy 93 AAGAACATTCCTTCTCTGATTGCAATTCGATTTGTTCTGATTCAGACATTCGCGACCA 152
Db 61 AAGAACATTCCTTCTCTGATTGCAATTCGATTTGTTCTGATTCAGACATTCGCGACCA 120
Qy 153 TCTTGCCTCGATTTGGTCAAAAGTCCAGAGTAAATACAGTGTGTTTACCTGTGCTT 212
Db 121 TCTTGCCTCGATTTGGTCAAAAGTCCAGAGTAAATACAGTGTGTTTACCTGTGCTT 180
Qy 213 AGTCCCAAAATATCTTCTGATTATATCCAGACTCGAAATTTTACAAGTGGAGCAAT 272
Db 181 AGTCCCAAAATATCTTCTGATTATATCCAGACTCGAAATTTTACAAGTGGAGCAAT 240
Qy 273 GTCAGACCATGAGAAATCCAGAAAGTTCATCGGCTTTACTGAAATCGGATTCGAGT 332
Db 241 GTCAGACCATGAGAAATCCAGAAAGTTCATCGGCTTTACTGAAATCGGATTCGAGT 300
Qy 333 GTTACTGTTCTGATGTGAGAGGTTTGGTGCACAGAGGTTTCTACCGAGAGACACGGT 392
Db 301 GTTACTGTTCTGATGTGAGAGGTTTGGTGCACAGAGGTTTCTACCGAGAGACACGGT 360
Qy 393 GCCTCTGAGTTCTCGGAGACAAATTTGTTCTAAAGTTAAGATCGAAATCGTTGTTAAG 452
Db 361 GCCTCTGAGTTCTCGGAGACAAATTTGTTCTAAAGTTAAGATCGAAATCGTTGTTAAG 420
Qy 453 AAAGACCAAGTGGATCTGTAATCAACAAATAATTGAAGGAGCAAGGACGAGAGATT 512
Db 421 AAAGACCAAGTGGATCTGTAATCAACAAATAATTGAAGGAGCAAGGACGAGAGATT 480
Qy 513 GGTGATGCAAGATTTTGTGTTGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 572
Db 481 GGTGATGCAAGATTTTGTGTTGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 540
Qy 573 CGTGGGGAGAGACGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 620
Db 541 CGTGGGGAGAGAGACGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 588
RESULT 6
AR125590
LOCUS
DEFINITION
AR125590
SEQUENCE 14 from patent US 6177275.
897 bp DNA linear PAT 16-MAY-2001
AR125590.1 GI:14111652
KEYWORDS
SOURCE
ORGANISM
Unknown.
Unclassified.

[illegible]

rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.	REFERENCE	1 (bases 1 to 796)	120185 bp	DNA	linear	PLN 01-APR-1999
Garcia-Ibáñeta, D. and Sengupta-Gopalan, C.	AUTHORS					
Characterization of PII (GLNB) in alfalfa	TITLE					
Unpublished	JOURNAL					
2 (bases 1 to 796)	REFERENCE					
Garcia-Ibáñeta, D. and Sengupta-Gopalan, C.	AUTHORS					
Direct Submission	TITLE					
Submitted (22-FEB-2001)	JOURNAL					
University, Corner of Knox and College, Las Cruces, NM 88003, USA	FEATURES					
Location/Qualifiers	source					
1..796	organism="Medicago sativa"					
/mol_type="mRNA"						
/db_xref="taxon:3879"						
1..796	gene					
/gene="GLNB"						
1..595	CDS					
/gene="GLNB"						
/codon_start=1						
/product="PII protein"						
/protein_id="AAK16221.1"						
/db_xref="GI:13277515"						
/translation="MTLIAKPNVFNGLNFNINETQIPFSSFSVIRKFGDSHRRNVYL RSGNASVLPKIRAOINLPDYVPSEKFKVEAILRPWIRIPOVSSGLLXMGKRGVTVS KFGAQGGSEFSESDNFVAKVMEIVRKDOVEAVINKIMETARTGEIGDK IFLIPVSDVIRIETGERGEQARMMAGLITDALYV"						
ORIGIN						
Query Match 31.1%; Score 253.8; DB 8; Length 796;						
Best Local Similarity 71.0%; Pred. No. 7.3e-50;						
Matches 336; Conservative 0; Mismatches 137; Indels 0; Gaps 0;						
QY 198 GTTTTACCTGCTGTAGTGGCCAAATATCTTCTGATTAATATTCAGACTCGAAATTTTAC 257						
Db 154 GTTCTTCCAAAATTAGAGCTCAAAACCTTCTGACTATGTCTCCGAATCCAAATTTAC 213						
QY 258 AAGTGGGAAGCAATTTGTGAGACCATGAGAAATCCAGCAAGTTTCATCGGCTTTACTGAAA 317						
Db 214 AAAGTTGAAGCAATTTCTCAGGCCATGAGAAATCCCTCAGTTTCTTCGGGTTTGTGAAA 273						
QY 318 ATCGGGAATTCGAGGTGTTACTGTTTCTGATGTGAGAGGTTTGTGTCACAAGAGGTTCT 377						
Db 274 ATGGAATTCGTTGTTGTTCTACTGATCTGATGTCAGGGGTTTGGTGCTCAGGGTGGCTCA 333						
QY 378 ACCGAGAGACACGGTGGCTCTGAGTTCTCGAAGACAAATTTGTGCTAAAGTTAAGATG 437						
Db 334 AAAGAGAGCGGAGGAGGCTCTGAAATTTCTGAAGACAATTTTGTGCCAAAGTTAAATG 393						
QY 438 GAATTCGTTGTTTAAAGAAAGACCAAGTGGAAATCTGTAATCAACACAATTAATTGAAGAGCA 497						
Db 394 GAATAGTGTGTGAGAAAAGACCCAGGTTGAGGAGTGAATAACAAAATTTATGGAGACGCA 453						
QY 498 AGGACAGAGAGATTCGTCATGGCAAGATTTTGTTCCTGCTGTGTAGATGTCATAGA 557						
Db 454 AGAACTGGGAGATTTGGTGTATGGCAAAAATTTTCTTGATCCCTGTATCTGATGTAATAAGA 513						
QY 558 GTTAGACAGGTGACGCTGGGAGAAAGCAGAGAGATGACTGGTGATATGCTTTTCAACCG 617						
Db 514 ATCCGACAGGTGAGCGTGGGAGCAGGCTGAGAGGATGGCTGGGGGACTAACTGACGCG 573						
QY 618 TCTTAGGAACAAACAGAGCTCAAGATGTTTTTTTTTTTTTTTTCATTTCCGTTCT 670						
Db 574 TTGATGTTTGAACGAAGATTAGCATTTGCAATTTGGTCTCTCTCTCTCTTCTT 626						
AC007138	RESULT 10					
Arabidopsis thaliana BAC T7B11 from chromosome IV near 10 cM,	LOCUS					
complete sequence.	DEFINITION					
AC007138	ACCESSION					

```

VERSION      AC007138.1  GI:4510323
KEYWORDS     HTG.
SOURCE       Arabidopsis thaliana (thale cress)
ORGANISM     Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE    1 (bases 1 to 120185)
AUTHORS      Huang, E.N., Nascimento, L., de la Bastide, M., Habermann, K.,
              Vil, M.D., Preston, R.R., Spiegel, L.A., See, L.H., Shah, R., Matero, A.,
              O'Shaughnessy, A., Rodriguez, M., Shekher, M., Swaby, I., Schutz, K.,
              Parnell, L.D., Dedhia, N.N. and McCombie, W.R.
              Arabidopsis thaliana BAC T7B11 from chromosome IV near 10 cM
              Unpublished
              2 (bases 1 to 120185)
              Huang, E.N., Nascimento, L., de la Bastide, M., Habermann, K.,
              Vil, M.D., Preston, R.R., Spiegel, L.A., See, L.H., Shah, R., Matero, A.,
              O'Shaughnessy, A., Rodriguez, M., Shekher, M., Swaby, I., Schutz, K.,
              Parnell, L.D., Dedhia, N.N. and McCombie, W.R.
              Direct Submission
              Submitted (23-MAR-1999) Lita Annenberg Hazen Genome Sequencing
              Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring
              Harbor, NY 11724
              3 (bases 1 to 120185)
              Parnell, L.D.
              Direct Submission
              Submitted (01-APR-1999) Lita Annenberg Hazen Genome Sequencing
              Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring
              Harbor, NY 11724
REMARK       Arabidopsis thaliana BAC T7B11 from chromosome IV near 9.5 cM
COMMENT      The T7 end of T7B11 is oriented toward the NOR and overlaps with
              T15B16, GenBank accession number AF104919. The SP6 end of T7B11 is
              oriented toward the centromere and overlaps with T10M13, GenBank
              accession number AF001308. There are no discrepancies found in
              these independently finished overlaps. The region from position
              110064 to 110253 exists as single stranded and single chemistry.
              Please see the feature below for details.
FEATURES     Location/Qualifiers
              1..120185
                 /organism="Arabidopsis thaliana"
                 /mol_type="genomic DNA"
                 /cultivar="Columbia"
                 /db_xref="taxon:3702"
                 /chromosome="IV"
                 /map="near 10 cM"
              misc_feature
                 1..14923
                    /note="overlap with T15B16, GenBank accession number
                    AF104919, from position 1 to 14923"
              50..1880
                 /gene="T7B11.1"
                 /note="encodes hypothetical protein; identical to
                 T15B16.9, GenBank accession number AF104919; gene model
                 last edited on 3 Mar 99"
                 /evidence=not experimental
                 join(50..556,1098..1259,1340..1561,1668..1880)
                 /gene="T7B11.1"
                 /note="identical to T15B16.9, GenBank accession number
                 AF104919; similar to T7B11.3; functional catalog ID=99"
                 /codon_start=1
                 /evidence=not experimental
                 /product="hypothetical protein"
                 /protein_id="AAD22637.1"
                 /db_xref="GI:4558544"
                 /translation="MAQKQTLHQRPFFSSPSRYSISNRPIFLSRNGLLVLAL
                 FILLGVLPWPGSPILLFPNKYSASLSHAKSEWRNTLAQAARFVATNGTVI
                 VCAVSSQFLPFLNNMLISVRKQKQKVLIAEDYITLYKVKMKPFGHVLIPALDLS
                 KTAYSQSGQFLNFTARRFQHLQLILELGNVNDVMDVWLQDPFQYLEGSHDAVFT
                 DMPQIKPLNHSDDPAPQNGETIYICSMIYLRPTNGAKLMMKWSBELQSOAWSES
                 TRFKANDQAFNALNAHQVDLYLLSQVAFPTGGLYFNDAARVKTGKHVIVNHN
                 YTIIGDRKMRFDQYGLWLVDHDALESPLGKLE"
                 complement(2647..4650)
                 /genes="T7B11.2"
                 /note="encodes putative CHP-rich zinc finger protein;
                 RKRGM"
              gene
                 /gene="T7B11.1"
                 /note="encodes hypothetical protein; identical to
                 T15B16.9, GenBank accession number AF104919; gene model
                 last edited on 3 Mar 99"
                 /evidence=not experimental
                 join(10826..10842,10929..11371,11559..11770,11843..12130,
                 12260..12433,12526..12762)
                 /gene="T7B11.4"
                 /note="identical to T15B16.7, GenBank accession number
                 AF104919; functional catalog ID=99"
                 /codon_start=1
                 /evidence=not experimental
                 /product="hypothetical protein"
                 /protein_id="AAD22640.1"
                 /db_xref="GI:4558547"
                 /translation="MGTRHCYGLGYEMFNFWSSSVRFVWSSSETLTKVITLRSIC
                 ILGKHTLLTFVTLFLKLELCQVSDCVIAETGLIVNIPPTKAQDGRPAGESGSK
                 LRDEYTLGRNNGLLFDKAYVDHGKMKMMKDGPKLGLYGIAGDDVNDGREG
                 ESAKNAIOERTWRIQKTVDDHERITKLLESEKXLEIKGNELAKPQVNGTMRKL
                 SEDLEORQKEELHEKILRLEROIDQVAIELEVLQKGLNWKMSBELQSOAWSES
                 DIIFKDLVEKEALADLNKFNQIILIRERTNDELOEARVELNMCSELVRKLPVDM
                 QQYICQEDRAVEVQLQWHERYINDFPHYKRVKLENDQREVEVIDDRDEKLRLK
                 ADIGDGPYNATVKALLLEINYPNSGYITTELWNFKEDKATLEEGVTCLLDQWEKAK
                 RKRGM"

```

gene	13384..14244 /gene="T7B11.5" /note="encodes RNaseP-associated protein; identical to T15B16.20, GenBank accession number AF104919; gene model last edited on 3 Mar 99" /evidence=not experimental Join(13384..13454,13566..13613,13697..13757,13842..13940,14020..14244) /gene="T7B11.5" /note="identical to T15B16.20, GenBank accession number AF104919; functional catalog ID=98" /codon_start=1 /evidence=not experimental /product="putative RNaseP-associated protein" /protein_id="AAD22642.1" /db_xref="GI:4558548" /translation="NRGVTPINPQKKIEADSNPVKESDYEGERLTHLDLIGRGT SXLNVNSPEKILWKKQIAIGINEVTRVLERMNPNTSDQQNPKQLQVTVIVADCK PMLTKHINLAASRVVLYVDNRKASRLGELVKLTALTAIGIKARGNDNLILK QILPRDS" complement(14442..19647) /genes="T7B11.6" /note="encodes putative SecA-type chloroplast protein transport factor; gene model last edited on 26 Mar 99" /evidence=not experimental complement(join(14442..14579,14663..14902,14991..15108,15196..15263,15359..15610,15686..15787,15888..15946,16035..16134,16215..16361,16449..16515,16710..16885,17107..17274,17533..17646,17718..17831,17987..18205,18310..18488,18577..18684,18783..18989,19082..19154,19231..19647)) /genes="T7B11.6" /note="functional catalog ID=06.04; functional catalog ID=30.26" /codon_start=1 /evidence=not experimental /product="putative SecA-type chloroplast protein transport factor" /protein_id="AAD22642.1" /db_xref="GI:4558548" /translations="MVSPLCDSSQLLYHRPISPTASQFVIADGIIHLRQNLSSSFW GTKFGNTVKGVSQSSCRKSTSVNAGLGLSGIFKSGDNGSTEQVYASIVASV NRLETSALSDSELRERTALKORAKGSMDSLLPEARAVREASRVLGLRPFDV QILGGMVHGEIAQMTGKTLVALPAYLNAISCKGVHVTVDYLARRCEWVG QVPELGLKVLQNMTPORKENYLDITITVNSLSEGDYLDNLATESVELVLR DFNYCVIDEVDSILIDEARTPLIISGPAEKPSQIYKAAKIASAFERDITHTDEKQK TVLLTEQGYEDABEILDVXDLYDPREQVASYVLNAIKELFLRDVNYIIRAKEVLIV DEFTGRVMOGRNSDGLHQAVEAKGLPIQNESITLASISYQNFPLQPKLCMTGTA STESAFESLYKLVTIVPTNKMIRKDESDVVFKAQVWVVEISRMHKTGRAV LVGTTSEQSDLSQLREAGITHEVLNAPENVEREAEIVASQSGRLGAVTIATNMAG
CDS	
gene	
FEATURES	source 1..198220 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /variety="Columbia" /db_xref="taxon:3702" /chromosome="4" 5603..8671 /gene="AT4g01810" complement(join(5603..5740,6062..6178,6259..8079,8177..8671)) /gene="AT4g01810" complement(join(5603..5740,6062..6178,6259..8079,8177..8671)) /gene="AT4g01810" /note="similarity to contains EST gb:H36218, AA651596, T42768, AI997386.1, AA395161" /codon_start=1 /product="putative protein transport factor" /protein_id="CAB80674.1" /db_xref="GI:7268565" /db_xref="GOA:Q9SY11" /db_xref="SPTREMBL:Q9SY11" /translation="MANLPKSSVNVVGTLTLEPNRSPQDRTPVHSPVPVAPIP PRFPQSPRQMSQSPSKSLLSPANGITGSPIPRLSTPGPPVPTVPKPAVP FRTPATPQPMAYSSANSLPVPSTPSFYNSGVSQSDLDLDVPMESPIRADSPVL FSAVKLVKQILANVAISGPAISVAGREISPGPOLIORDHRCNLCCAYSNPYSIL IGSGQMCCIENMNGSKGRYVASSKNELQNPFLSLVDVDTQGNKRGVFPASDS RTSAPVLIIDECLDPEHLQSLHAFVDSLPQTLTGILTYGRVTLYNVPASRD ASADVISAKSPSABSMKALYGTGVYLSPMHSLKVAHEIFSSRLTYGPGVHMSHPN RCLGTVAEALAIQGPSAERGVRVAGNSRIIVCAGGPIYKPGSPVHMSHPN YPYMEKTAIKWENIGRAHRNTVVDILCAGTCILRPILQPLAKSGVLVLDHDF GEAFVDLQRAATRAAGSHGLLEVRCSDDILITQVIGPEEAHSETHETFKSDAALSI QMLSVETQSFSLSMENKRDIKSDHVFQFAFYSDVYQADVSRVTFKLPTVLSIA YLQSVEDRASVALISKRTLLAKNOKADMEATVDRIKDIALFKFGSPKSKLYSF PKELSLPELLFLHREPLNIGIHEDERSVLRNLFLNLSFDLSLRVAPRCLMHQE GGTPEELPAYDLSMQSDKAVILDHGTDFIWLGAELSADEVKSAVLAACRTLABELT


```

Db 52225 TAAATATGAATGAGTCTGTGTTTGGTTTCCTGTTTGAATGATCAAGATGTTTTCAC 52284
QY 745 TGTACATGAATATGACAGAAACATCTGCTCTGTTCTCAGACATCGAAATCTGTGTCCTA 804
Db 52285 TGTACATGAATATGACAGAAACATCTGCTCTGTTCTCAGACATCGAAATCTGTGTCCTA 52344
QY 805 ATAAAA 810
Db 52345 ATAAAA 52350

RESULT 12
AY442185 947 bp mRNA linear PLN 15-NOV-2003
LOCUS Lycopersicon esculentum PII-like protein (GLBI) mRNA, complete cds.
ACCESSION AY442185
VERSION AY442185.1 GI:38231569
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
REFERENCE Walch-Liu, P., Roembeld, V. and von Wiren, N.
AUTHORS Walch-Liu, P., Roembeld, V. and von Wiren, N.
TITLE A PII-like protein in tomato
JOURNAL Unpublished
REFERENCE Walch-Liu, P., Roembeld, V. and von Wiren, N.
AUTHORS Walch-Liu, P., Roembeld, V. and von Wiren, N.
TITLE Direct Submission
JOURNAL Submitted (17-OCT-2003) Institute of Plant Nutrition, Hohenheim University, Fruwirthstr.20, Stuttgart 70593, Germany
FEATURES
    source      1..947
                /organism="Lycopersicon esculentum"
                /mol_type="mRNA"
                /db_xref="taxon:4081"
    gene        1..947
                /gene="GLBI"
                /note="LeGLB1"
    CDS          53..652
                /gene="GLBI"
                /note="similar to nitrogen regulatory protein PII"
                /product="PII-like protein"
                /protein_id="AA14689.1"
                /db_xref="GI:38231570"
                /translation="MASPSLSKNSFSLHSFSSPSLSQFPHTSITVVPKPPFSOLT
                FKCONAPSEPIIAQNSPDFDPAKYVEALIPWRIQOVSSALLKMGIRGVTVSD
                VRGFGAGGLTERQAGSEFSDTFVAKYMEIVSKQVEGVIAKIEARTGELGDC
                KIFLTPISDVIRVTRGERGEKAERMGMGHADYSSALSTS"
ORIGIN
    Query Match 30.0%; Score 245; DB 8; Length 947;
    Best Local Similarity 72.5%; Pred. No. 8.9e-48;
    Matches 317; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 167 GGTCAAAAGTACCGAGTAATAACAGTCGTGTTTACCTGCTGTTAGTGCCTCAATATC 226
Db 178 GCTCACCITTCAAACGCTGTCAAAATGCTCTCTTCCGATTATCAGAGCCCAAAATC 237
QY 227 TTCTGATTATATTCAGACTCGAAATTTTACAGGTGGAGCAATGTTCAGACCATGAG 286
Db 238 TCCAGATTTCGTCCCTGTATCGGAAGTTTACAAAGTTGAAGCAATTTCTAAGACCTTGGAG 297
QY 287 AATCCAGAGTTTCATCGCTTTTACTGAAATCGGATTTCGAGGTGTTACTGTTCTGA 346
Db 298 AATTCAACAGGTTTCTTCGGCACTACTGAAATGGGCATTCGTGTGTCTACGTTTCGA 357
QY 347 TGTGAGAGGTTTGTGTCACAAAGGAGTTTCTACCGAGACACAGGTGGCTCTGAGTTCTC 406
Db 358 TGTTCTGTTTGGCGCCCAAGTGGCTTGACTGAGAGGCAAGCTGGCTCTGAAATCTTC 417

```

```

QY 407 GGAAGACAAATTTGTTGCTAAAGTTAAGATGGAATTCGTTTGAAGAAAGACCAAGTGA 466
Db 418 TGAAGACACGTTTGTGCAAAAGTTAAATGGAATTTGTTGTCAGCAAGACCAAGTGA 477
QY 467 ATCTGTAATCAACACAATAATTGAAGGAGCAAGACAGGAGATTGTTGTCAGCAAGAT 526
Db 478 AGGAGTCATTGCCAAGATAATTGAAGAGGCAAGAACTGGTGAATAGGTGATGGAAGAT 537
QY 527 TTTTGTGTTTGCCTGTGTCAGATGTCATAAGAGTTTAGCACAGGTGAGCGTGGAGAAAGC 586
Db 538 ATTCTTGACTCCCATCTCCGATGTTATAAGAGTTCGCACCTGGTGAACGGGGAGAAAGCC 597
QY 587 AGAGACATGACTGGTG 603
Db 598 AGAGAGATGATGGAG 614

RESULT 13
PPI489604 1085 bp mRNA linear PLN 11-OCT-2003
LOCUS Pinus pinaster mRNA for PII-like protein (glb gene).
ACCESSION AJ489604
VERSION AJ489604.1 GI:37653226
KEYWORDS glb gene; PII-like protein.
SOURCE Pinus pinaster
ORGANISM Pinus pinaster
REFERENCE Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
TITLE Canton, F.J.
JOURNAL A PII-like protein from xylem of adult Maritime pine trees
REFERENCE 2 (bases 1 to 1085)
AUTHORS Canton, F.J.
TITLE Direct Submission
JOURNAL Submitted (11-JUN-2002) Canton F.J., Molecular Biology and Biochemistry, Faculty of Sciences, Campus de Teatinos s/n, E-29071, SPAIN
FEATURES
    source      1..1085
                /organism="Pinus pinaster"
                /mol_type="mRNA"
                /db_xref="taxon:71647"
                /tissue_type="xylem"
                /dev_stage="adult tree"
    gene        1..1085
                /gene="glb"
    CDS          42..773
                /gene="glb"
                /function="putative nitrogen sensor protein"
                /codon_start=1
                /product="PII-like protein"
                /protein_id="CAD33967.1"
                /db_xref="GI:37653227"
                /translation="MAHLPLFVKGSIFLSPSTSSCTVSSISHTAASPCFHSSP
                AAANKSVLGTQMQRMDGVRVSKRMKQSLHAKMEKQKQKAKAQAPNGTNTHPDYVPEA
                NFYKVEALIRPWRISHVTTGLLKMGIQVTVSDVRGFGVQAGSAERQAGSESKNFV
                SKYMEIVVSKQVEAVIDAIDEARTGEIGDKIFVVPVADVIRVTRGERDLKPERM
                AGGORYLQVYIKRIQTAKSLAL"
ORIGIN
    Query Match 25.7%; Score 209.6; DB 8; Length 1085;
    Best Local Similarity 70.9%; Pred. No. 2.3e-39;
    Matches 278; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 212 TAGTCCCAATATCTTCTGATTATATTCAGACTCGAAATTTTACAGGTGGAAGCAAT 271
Db 311 TGGGACCAACACTCATCCAGATTATGTCACAGAGCCAACTTTTACAAAGTAGAAGCAAT 370
QY 272 TGTGAGACATGAGATCCAGAGTTTCTCGGTTTACTGAAATCGGATTTCGAGG 331
Db 371 ATTGAGGCCATGGCGCATCTCCACTGTACTACGGGTCTATTGAAATGGGATTCTGG 430

```

TITLE
JOURNAL

COMMENT

Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yanada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan [E-mail: skkikuchien@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007]

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

PAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirao, T., Hori, F., Iida, J., Inamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Satoh, H., Sakai, K., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

FEATURES
source
 i. 902
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="J013149B08"

ORIGIN

Query Match 24.4%; Score 199.4; DB 8; Length 902;
Best Local Similarity 71.3%; Pred. No. 6.4e-37;
Matches 263; Conservative 0; Mismatches 106; Indels 0; Gaps 0

QY 240 CCAGACTCGAAATTTTACAGGCTGGAGCAATTTCTCAGACCATCGAGNATCCAGCACTT 299
DB 283 CCGGAGTCGGAGTTCTCAAGGTCGGAGCAATCTCGGCCCATCGAGGTCCTTATGTC 342
QY 300 TCATCGGCTTCTCTGAAATCGCGATTTCAGGCTGTACTGTTTCTGATGTCAGAGGGTTT 359
DB 343 TCATCGGCTTCTCTGAAATCGCGATTTCAGGCTGTACTGTTTCTGATGTCAGAGGGTTT 402
QY 360 GGTGTCAAGAGGCTTCTACCGAGAGACA CGTGGCTCTGAGTTCTCGAAGACAAATTT 419
DB 403 GCGCCACAGGCGGCTCACTGAGAGGCATCAGGGTCAGAAATTTGCAAGAGATACATT 462
QY 420 GTTGCTAAAGTTAAGATCGAAATCGTTGTTTAAAGAAAGCCAGGTGAATCTGTAATCAAC 479
DB 463 ATTCATAAAGTTAAGATCGAAATAGTGTGTGCCAGGATCAGGTTGAAGCTGTGTTGTCAC 522
QY 480 ACHATAATTGAGAGCAGAGCAGGAGGAGTTGGTGATGCCAAGATTTTGTGTTGCT 539
DB 523 AAGATAAATTGAAAGGCAAGAACCGAGGAAATTTGGTGATGGAATAATATTTTGATACC 582

Search completed: May 27, 2004, 22:22:39
Job time : 2419.14 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2004, 13:28:59 ; Search time 284.934 Seconds
(without alignments)
12180.972 Million cell updates/sec

Title: US-09-756-541-13

Perfect score: 817

Sequence: 1 CTGAAGTGTGTAAATAA.....GTTCTAATAAATAAATAA 817

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001as:*
- 5: Geneseq2001bs:*
- 6: Geneseq2002s:*
- 7: Geneseq2003as:*
- 8: Geneseq2003bs:*
- 9: Geneseq2003cs:*
- 10: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	817	100.0	817	4	Aaf58581 Arabidops
2	591	72.3	591	6	Abz14328 Arabidops
3	588	72.0	588	4	Aaf58583 Arabidops
4	286.4	35.1	897	4	Aaf58582 Ricinus c
5	285	34.9	594	4	Aaf58584 Ricinus c
6	92.8	11.4	96109	4	Aaf28548 Genomic f
7	85.8	10.5	375	8	Ada30024 DNA encod
8	84.6	10.4	110000	2	Continuation (2 of
9	84.4	10.3	110000	2	Continuation (15 of
10	81.6	10.0	110000	7	Continuation (4 of
11	81.6	10.0	110000	7	Continuation (4 of
12	81.4	10.0	339	7	Acf72027 Photorhab
13	79.8	9.8	110000	2	Continuation (4 of
14	77	9.4	363	7	Abz40022 N. gonorr
15	73	8.9	16526	3	Aaa81472 N. mening
16	73	8.9	110000	3	Continuation (7 of
17	73	8.9	172325	3	Aaf21613 Neisseria
18	73	8.9	349980	3	Aaf21612 Neisseria
19	68.8	8.4	336	5	Aah68445 C glutami
20	68.8	8.4	417	5	Aah67222 C glutami
21	68.8	8.4	349980	5	Aah68531 C glutami
22	67.2	8.2	1164	5	Aas88441 DNA encod
23	67.2	8.2	11095	4	Aas46256 DNA encod

24	64.6	7.9	663	7	ACF65588	Acf65588 Photorhab
25	64.6	7.9	110000	7	ACF67367_46	Continuation (47 o
26	64.6	7.9	110000	7	ACF65388_01	Continuation (2 of
27	61	7.5	339	7	ACF71482	Acf71482 Photorhab
28	60.4	7.4	110000	6	ABQ69245_15	Continuation (16 o
29	60.4	7.4	110000	6	ABQ67195_1	Continuation (2 of
30	60.4	7.4	110000	6	ABQ67195_2	Continuation (3 of
31	59.6	7.3	290	4	Aaf71296	Aaf71296 Corynebac
32	57.6	7.1	110000	6	ABA90521_16	Continuation (17 o
33	52.4	6.4	2751	6	ABQ70750	Abq70750 Listeria
34	50.8	6.2	110000	6	ABA03041_15	Continuation (16 o
35	49.6	6.1	110000	4	AAI99682_32	Continuation (33 o
36	49.6	6.1	110000	4	AAI99683_32	Continuation (33 o
37	47	5.8	342	6	ABN18985	Abn18985 Human ORF
38	45.8	5.6	9733	6	ABL32682	Ab132682 Human imm
39	44.4	5.4	174	7	ABZ39932	Abz39932 N. gonorr
40	43.6	5.3	2000	7	ADA71938	Ada71938 Rice gene
41	43.4	5.3	3858	5	AAS89065	Aas89065 DNA encod
42	43.4	5.3	3858	5	AAS90406	Aas90406 DNA encod
43	43.4	5.3	3858	5	AAS85755	Aas85755 DNA encod
44	43.4	5.3	3858	5	AAS92604	Aas92604 DNA encod
45	42.8	5.2	1970	3	AAZ99179	Aaz99179 T. hydrot

ALIGNMENTS

RESULT 1
AAF58581
ID AAF58581 standard; cDNA; 817 BP.
XX
AC AAF58581;
XX
DT 23-APR-2001 (first entry)
XX
DE Arabidopsis thaliana P-PII cDNA.
XX
KW Arabidopsis thaliana; PII; plant nitrogen regulatory gene; P-PII;
KW nitrogen assimilation; transgenic plant; herbicide screening; ss.
XX
OS Arabidopsis thaliana.
XX
PN US6177275-B1.
XX
PD 23-JAN-2001.
XX
PF 23-JUL-1997; 97US-00899330.
XX
PR 24-JUL-1996; 96US-0022328P.
XX
PA (UNYNY } UNIV NEW YORK STATE.
XX
PI Coruzzi GM, Lam H, Hsieh M;
XX
DR WPI; 2001-158572/16.
DR P-PSDB; AAB69495.
XX
PT Novel P-PII genes capable of regulating plant nitrogen assimilation,
PT useful for transgenic plant production, and as probes for isolating
PT additional genomic clones having P-PII gene promoters.
XX
PS Claim 2; Fig 12; 35pp; English.
XX
CC The present sequence encodes a nitrogen regulatory PII protein. Novel
CC plant PII (also called P-PII) nucleotide sequences have been isolated.
CC They are useful for regulating nitrogen assimilation in plants, and in
CC transgenic plant production. They are also used to engineer organisms
CC that overexpress wild-type or mutant P-PII regulatory proteins. P-PII
CC proteins are useful for in vitro screening of herbicides. P-PII
CC nucleotides may be used as probes for isolating additional genomic clones
CC with the promoters of P-PII genes. P-PII promoters are light- and/or
CC sucrose-inducible, and are suitable for genetic engineering of plants

SQ Sequence 817 BP; 242 A; 137 C; 188 G; 250 T; 0 U; 0 Other;
Query Match 100.0%; Score 817; DB 4; Length 817;
Best Local Similarity 100.0%; Pred. No. 3.8e-220;
Matches 817; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTGAAGTGTGTTAAATAAATACTAGATCATGGGGGCTCAATGACGAAACCCATCT 60
DB 1 CTGAAGTGTGTTAAATAAATACTAGATCATGGGGGCTCAATGACGAAACCCATCT 60
QY 61 CAATAACTTCTCGGTTCTATCTCATGCAAGAACATGCTTCTCTGATTGCAATTT 120
DB 61 CAATAACTTCTCGGTTCTATCTCATGCAAGAACATGCTTCTCTGATTGCAATTT 120
QY 121 CGATTGTTCTGGATTGAGCAATCCCGACCATCTGCTCGATTGGTCAAAAGTCAC 180
DB 121 CGATTGTTCTGGATTGAGCAATCCCGACCATCTGCTCGATTGGTCAAAAGTCAC 180
QY 181 CGAGTAATAACAGTCGTTTAACTGTCGTTAGTGGCCCAATATCTCTGATTATATC 240
DB 181 CGAGTAATAACAGTCGTTTAACTGTCGTTAGTGGCCCAATATCTCTGATTATATC 240
QY 241 CAGACTCGAATTTTACAGGTGGAAGCAATTTGTCAGACCATGAGAAATCCAGCAAGTTT 300
DB 241 CAGACTCGAATTTTACAGGTGGAAGCAATTTGTCAGACCATGAGAAATCCAGCAAGTTT 300
QY 301 CATCGGCTTTACTGAAATCGGGATTGAGGTGTTTCTGTTTCTGATGAGAGGTTTG 360
DB 301 CATCGGCTTTACTGAAATCGGGATTGAGGTGTTTCTGTTTCTGATGAGAGGTTTG 360
QY 361 GTGCACAGGAGGTTTACCGAGAGACACCGTGGCTGAGGTTCTCGGAACAAATTTG 420
DB 361 GTGCACAGGAGGTTTACCGAGAGACACCGTGGCTGAGGTTCTCGGAACAAATTTG 420
QY 421 TTGCTAAAGTTAAGATGAAATCGTTGTTTAAAGAACCAAGTGAATCTGTAATCAACA 480
DB 421 TTGCTAAAGTTAAGATGAAATCGTTGTTTAAAGAACCAAGTGAATCTGTAATCAACA 480
QY 481 CAATAATTTGAAGGAGCAGGACAGAGAGATTGGTGTGATGGCAAGATTTTGTGCTG 540
DB 481 CAATAATTTGAAGGAGCAGGACAGAGAGATTGGTGTGATGGCAAGATTTTGTGCTG 540
QY 541 TGTGATGTCATGAGTGTAGGACAGGTGAGCGTGGGAGAACGACAGAGATGACTG 600
DB 541 TGTGATGTCATGAGTGTAGGACAGGTGAGCGTGGGAGAACGACAGAGATGACTG 600
QY 601 GTGATATGCTTTTACCGTCTTAGGACAAACAGAGCTCAAGATCGTTTCTTTTCTTTC 660
DB 601 GTGATATGCTTTTACCGTCTTAGGACAAACAGAGCTCAAGATCGTTTCTTTTCTTTC 660
QY 661 ATTTGCGTCTCTAGATTCTGCGAATAAATGAAATGAGTCTGTTGTTGTTTCAATGTT 720
DB 661 ATTTGCGTCTCTAGATTCTGCGAATAAATGAAATGAGTCTGTTGTTGTTTCAATGTT 720
QY 721 AATCGATCAAGATGTTTAACTGTACATCAATATGTCAGAAACATCTGCTGCTGTC 780
DB 721 AATCGATCAAGATGTTTAACTGTACATCAATATGTCAGAAACATCTGCTGCTGTC 780
QY 781 TCAGACATCGAACTCTGTTCTCTAATAAATAAATAA 817
DB 781 TCAGACATCGAACTCTGTTCTCTAATAAATAAATAA 817

RESULT 2
AB214328
ID AB214328 standard; DNA; 591 BP.
XX AC AB214328;
XX AC
XX 21-JAN-2003 (first entry)
XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 2133.

XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX Arabidopsis thaliana.
XX WO200216655-A2.
XX 28-FEB-2002.
XX 24-AUG-2001; 2001WO-US026685.
XX 24-AUG-2000; 2000US-0227866P.
XX 26-JAN-2001; 2001US-0264647P.
XX 22-JUN-2001; 2001US-0300111P.
XX (SCRI) SCRIPPS RES INST.
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX Harper JF, Kreps J, Wang X, Zhu T;
XX WPI; 2002-304127/34.
XX Identifying a stress condition to which a plant cell has been exposed and
XX producing plants with increased tolerance to these abiotic stresses.
XX Claim 144; SEQ ID NO 2133; 577pp + Sequence Listing; English.
XX The invention relates to identifying a stress condition to which a plant
XX cell has been exposed, comprising: (a) contacting nucleic acid
XX representative of expressed polynucleotides in the plant cell with an
XX array or probes representative of the plant cell genome; and (b)
XX detecting a profile of expressed polynucleotides in the plant cell
XX characteristic of a stress response. The method is useful in the
XX production of transgenic plants, cells and seeds and in producing plants
XX with increased tolerance to abiotic stress. The present sequence is that
XX of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
XX in methods of the invention. Note: The sequence data for this patent is
XX not represented in the printed specification but is based on sequence
XX information supplied to Derwent by the European Patent Office
XX
XX Sequence 591 BP; 165 A; 105 C; 148 G; 173 T; 0 U; 0 Other;
Query Match 72.3%; Score 591; DB 6; Length 591;
Best Local Similarity 100.0%; Pred. No. 2e-156;
Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 33 ATGGCGGCTCAATGACGAAACCCATCTCAATACTTCTCTCGGTTCTATTTCTGATCGA 92
DB 1 ATGGCGGCTCAATGACGAAACCCATCTCAATACTTCTCTCGGTTCTATTTCTGATCGA 60
QY 93 AAGACATTGCTTCTCTGATTGCAATTTCTGTTCTGATTGCAATTTCCCGACCA 152
DB 61 AAGACATTGCTTCTCTGATTGCAATTTCTGTTCTGATTGCAATTTCCCGACCA 120
QY 153 TCTTGCCTCGATTTGGTCAAAAAGTCACCGAGTAATAACAGTCGTTGTTTACTGCTGTT 212
DB 121 TCTTGCCTCGATTTGGTCAAAAAGTCACCGAGTAATAACAGTCGTTGTTTACTGCTGTT 180
QY 213 AGTGCCAAATATCTTCTGATTATATTCAGACTCGAAATTTTACAAGTGAAGCAAT 272
DB 181 AGTGCCAAATATCTTCTGATTATATTCAGACTCGAAATTTTACAAGTGAAGCAAT 240
QY 273 GTCAGACCATGAGAAATCCAGCAAGTTTCTATCGGCTTTTACTGAAAATCGGATTCGAGT 332
DB 241 GTCAGACCATGAGAAATCCAGCAAGTTTCTATCGGCTTTTACTGAAAATCGGATTCGAGT 300
QY 333 GTTACTGTTTCTGATGTGAGAGGTTTGTGCAAGGAGGTTTCTACCGAGAGACACGTT 392
DB 301 GTTACTGTTTCTGATGTGAGAGGTTTGTGCAAGGAGGTTTCTACCGAGAGACACGTT 360
QY 393 GGCTCTGAGTTCTCGAAGACAAATTTGTTGCTAAAGTTTAAAGATGAAATCGTTGTTAAG 452
DB 361 GGCTCTGAGTTCTCGAAGACAAATTTGTTGCTAAAGTTTAAAGATGAAATCGTTGTTAAG 420

CC that overexpress wild-type or mutant P-II regulatory proteins. P-II
 CC proteins are useful for in vitro screening of herbicides. P-II
 CC nucleotides may be used as probes for isolating additional genomic clones
 CC with the promoters of P-II genes. P-II promoters are light- and/or
 CC sucrose-inducible, and are suitable for genetic engineering of plants
 XX Sequence 397 BP; 254 A; 171 C; 218 G; 254 T; 0 U; 0 Other;

Query Match 35.1%; Score 286.4; DB 4; Length 397;
 Best Local Similarity 69.8%; Pred. No. 2.3e-70;
 Matches 402; Conservative 0; Mismatches 171; Indels 3; Gaps 1;
 QY 28 GAATCATGGCGGCTCAATGACGAACCCCATCTCAATACCTCTCGGTTTCTATCTG 87
 DB 45 GAATTCGGCAGCGGCTACTGCGAACTGGGCTTGCTCACTCTCTCAATCTTAATAACA 104
 QY 88 ATCGAAGAACATTCCTTCTGATTCGATTTGATTTGCTTGGATTCCAGCATTCCTCC 147
 DB 105 TCAGAAAGAAATCCCTGTTTGAATTCAGTTTGTGTCAGAGCTTAGACATTCCTC 164
 QY 148 GACCATCTTGCCCTCGATTTGGTCAACAAAGTCACCGAGTAATAACAGTCGGTTTACCTG 207
 DB 165 GGTTCCTCACTTAACACCGCGGTCAAGCGGTAAAGATGCTCC--CCCGTGGTTCCTG 221
 QY 208 TCGTTAGTCCCAATATCTTCTGATTTATTCAGACTCGAAATTTTACAGTGGAG 267
 DB 222 TGATTAATGCCAAAGCTCCGCTGACTACATTCCTGATGCTAAATCTCAAAAGTGAAG 281
 QY 268 CAATTCGACACCATGAGAAATCAGCAAGTTTCATCGGCTTTTACTGAAATCGGATTC 327
 DB 282 CAATTCAGCCCTGGCGAGTCTCGCAAGTTTCTCGGCTTCTGCTAAATTTGGTATTC 341
 QY 328 GAGGTGTACTGTTCTGATGTGAGAGGTTTGGTGCACAAAGGAGTTCTACCGAGAGAC 387
 DB 342 GAGGTGTACTGTTCTGATGTTCGAGGTTTGGTGCCTCAAGTGGTTCAACTGAGAGGC 401
 QY 388 ACSGTGCTGAGTTCCTCGAAGCAAAATTTGTTGCTTAAAGTTTAAAGTGAATTCGTTG 447
 DB 402 AGGGCGGCTCAGAAATTTCTGAAACAAAGTTTGTGCTTAAAGTTTAAAGTGAATTCGTTG 461
 QY 448 TTAAGAAAGACCAAGTGGATCTGTAATCAACAAATAATTAAGGAGGAGGACAGGAG 507
 DB 462 TTAGCAAGACCAAGTTGAGGATGTATAGAAAATCATTGAGGAGGCAAGACTGGAG 521
 QY 508 AGATTGCTGATGCAAGATTTTGTGCTGCTGTCAGATGTCATTAAGTTAGGACAG 567
 DB 522 AGATTGAGACGCGACAGATTTCTTGCTGCTGTTTCAAGTGTAAAGAGTTCGCGCATG 581
 QY 568 GTGAGCGTGGGAGAAAGCAGAGAGATGACTGGTG 603
 DB 582 GTGAGCGGGGTGATAAGCTGAGAGGATGACAGGAG 617

RESULT 5
 AAF58584
 ID AAF58584 standard; cDNA; 594 BP.
 XX AC AAF58584;
 XX AC AAF58584;
 DT 23-APR-2001 (first entry)
 XX Ricinus communis P-II cDNA fragment.
 KW Castor bean; PII; plant nitrogen regulatory gene; P-II;
 KW nitrogen assimilation; transgenic plant; herbicide screening; ss.
 XX Ricinus communis.
 XX US6177275-B1.
 XX 23-JAN-2001.
 XX 23-JUL-1997; 97US-00899330.

XX 24-JUL-1996; 96US-0022328P.
 XX (UUNY) UNIV NEW YORK STATE.
 XX Coruzzi GM, Lam H, Hsieh M;
 XX WPI; 2001-158572/16.
 DR P-FSDB; AAB69496.
 XX Novel P-II genes capable of regulating plant nitrogen assimilation,
 PT useful for transgenic plant production, and as probes for isolating
 PT additional genomic clones having P-II gene promoters.
 XX Claim 2; Col 37-38; 35pp; English.
 XX The present sequence encodes a nitrogen regulatory PII protein. Novel
 CC plant PII (also called P-II) nucleotide sequences have been isolated.
 CC They are useful for regulating nitrogen assimilation in plants, and in
 CC transgenic plant production. They are also used to engineer organisms
 CC that overexpress wild-type or mutant P-II regulatory proteins. P-II
 CC proteins are useful for in vitro screening of herbicides. P-II
 CC nucleotides may be used as probes for isolating additional genomic clones
 CC with the promoters of P-II genes. P-II promoters are light- and/or
 CC sucrose-inducible, and are suitable for genetic engineering of plants
 XX Sequence 594 BP; 149 A; 117 C; 157 G; 171 T; 0 U; 0 Other;

Query Match 34.9%; Score 285; DB 4; Length 594;
 Best Local Similarity 70.3%; Pred. No. 4.7e-70;
 Matches 397; Conservative 0; Mismatches 165; Indels 3; Gaps 1;
 QY 39 GCGTCATGACGAAACCCATCTCAATAACTCTCTCGGTTTCTATCTGATCGAAAGAAC 98
 DB 7 GAGGCTACTCGAAACTGGGCTTGCTCACTCTCTTCAATTAATATCAATCAAGAAAGAA 66
 QY 99 ATTGCTTTCTCGATTCGATTTTCGATTTGTTCTGATTCAGACATTCGACCATTTTCG 158
 DB 67 TTCCCTGTTTGTGATTTTCAGTTTGTGTTGTCAGAGCTTAGACATTCGTTTTCTCAC 126
 QY 159 CTCGATTTGGTCAAAAGTCACCGAGTAATAACAGTCGTGTTTACCTGTGTTAGTGCC 218
 DB 127 TTTAAACACCGCGGTCAAGCGGTAAAGATATGCC---CCCGTCGTTCTGTGATTAATGCC 183
 QY 219 CAATATATCTTGATTAATATTCAGACTCCGAAATTTTACAGGTGGAGCAATTTGTCAGA 278
 DB 184 CAAAGCTCGCTGACTACTCAATTCCTGATGCTTAATTTCTACAAAGTGGAGCAATTTCTCAGG 243
 QY 279 CCATGGAAGATCCAGCAAGTTTCACTCGGCTTTTACTGAAAATCGGATTCGAGGTGTTACT 338
 DB 244 CCTGGCGAGTCTCGCAAGTTTCTCGGCTTTGCTTAAATTTGTTATTCGAGGTGTTACT 303
 QY 339 GTTCTGATGTGAGAGGTTTGGTGCACAGAGGTTTCTACCGAGAGACAGCGTGCTCT 398
 DB 304 GTTCTGATGTTCGAGGTTTGGTGTCTCAAGGTGGTTCACTGAGAGGACGCGCGGCTCA 363
 QY 399 GAGTTCGCGAGACAAATTTGTTGCTTAAAGTTAAGATGGAATCGTTTGAAGAAAGAC 458
 DB 364 GAATTTCTGAAGACAAAGTTTGTGCTTAAAGTTAAGATGGAATCGTTTGAAGAAAGAC 423
 QY 459 CAAGTGAATCTGTAATCAACAAATTAATGAGAGGACCAAGACAGGAGAGATTGGTGTAT 518
 DB 424 CAGGTTGAGGATGTTATAGAAAATTAATCATTGAGGAGGCAAGAACTGGAGAGATTGGAGAC 483
 QY 519 GCGAAGATTTTGTTCCTGTCATGATGTCATGATGTCATGATGTCATGATGTCATGATGTCG 578
 DB 484 GCGAAGATTTTGTTCCTGTCATGATGTCATGATGTCATGATGTCATGATGTCATGATGTCG 543
 QY 579 GAGAAAGCAGAGAGATGACTGGTG 603
 DB 544 GATAAGGCTGAGAGATGACAGGAG 568

```
RESULT 6
AAF28548
ID AAF28548 standard; DNA; 96109 BP.
XX
XX AAF28548;
AC
XX
XX 04-APR-2001 (first entry)
XX
XX Genomic fragment #35.
XX
XX Genomic library; bacteria; human upper airway; otitis media; sinusitis;
XX bronchopulmonary; endocarditis; meningitis; ss.
XX
XX Moraxella catarrhalis.
OS
XX WO200078968-A2.
XX
XX 28-DEC-2000.
XX
XX 16-JUN-2000; 2000WO-US016649.
XX
XX 18-JUN-1999; 99US-0140121P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Legace RE, Patterson C, Berg KL;
XX
XX WPI; 2001-041427/05.
XX
XX Genomic library for identifying diagnostic and therapeutic compositions,
XX and for identifying virulence factors, regulatory elements and drug
XX targets, comprises Moraxella catarrhalis nucleic acids.
XX
XX Claim 1; Page 345-368; 545pp; English.
XX
XX The present invention relates to a Moraxella catarrhalis genomic library
XX comprising of a combination of 41 nucleic acid molecules (see AAF28548-
XX AAF28548). The library has a number of uses described in the
XX specification e.g. is useful for identifying diagnostic and therapeutic
XX compositions. M. catarrhalis (Branhamella catarrhalis) is a large
XX aerobic, gram-negative diplococcus, normally found among the bacterial
XX flora of human upper airways. M. catarrhalis is known to cause acute,
XX localised infections such as otitis media, sinusitis and bronchopulmonary
XX infection and life-threatening, systemic diseases including endocarditis
XX and meningitis
XX
XX Sequence 96109 BP; 28783 A; 18910 C; 20341 G; 28075 T; 0 U; 0 Other;
XX
XX Query Match 11.4%; Score 92.8; DB 4; Length 96109;
XX Best Local Similarity 57.8%; Pred. No. 7.1e-15;
XX Matches 185; Conservative 0; Mismatches 132; Indels 3; Gaps 1;
XX
XX 267 GCAATTGTCAGACCATGGAGATCCAGCAAGTTTCATCGGCTTTTACTGAAATCGGGATT 326
XX
XX 12953 GCAATTATACAGCCGTTTAACTCGATGTCGAGAGCACTCTCAGAAATGCGGTC 13012
XX
XX 327 CGAGGTGTACGTTTTCGATGTCGAGAGGTTTGGTGCACAGGAGGTTCTACCGAGAGA 386
XX
XX 13013 AATGGTATACCGTCAGTCAAGTCAAGGCTTTGGTGCACAAAGGTCATACCGGATG 13072
XX
XX 387 CACGGTGGCTCTGAGTCTCGCAAGCAAAATTTGTTGCTAAAGTTAAGATGGAATCGTT 446
XX
XX 13073 TATCGTGGGCGGAATATGTTGTTGAT---TTTTCACAAATTAATTAATGAGATAGCA 13129
XX
XX 447 GTTAAGAAGAGCAAGTGGAAATCTGTAATCAACACAATTAATTAAGGAGCAAGGACAGA 506
XX
XX 13130 TGTCTGATGAGATGTTGATCAATTAATGAGTCAATCATTAAGTTCGAAATACAGGT 13189
XX
XX 507 GAGATGTGGTATGCGCAAGATTTTGTGTTTCCCTGTGTCTAGATGTCTAAGAGTTAGGACA 566
XX
XX 13190 AAAATGGTGTGATGGTAAGATTTTGTGTTAGTCCGCTTGAGCGGTGTCATTCGCAATTCGA 586
XX
XX 567 GGTGACGCTGGGAGAAAC 586
```

```
Db 13250 GGCGAATTGATGAAGTGC 13269
||| ||| ||| ||| |||
RESULT 7
ADA30024
ID ADA30024 standard; DNA; 375 BP.
XX
XX ADA30024;
AC
XX
XX 20-NOV-2003 (first entry)
XX
XX DNA encoding Acinetobacter baumannii protein #1311.
XX
XX ds; gene; Acinetobacter baumannii; bacterial disease; antibacterial;
XX vaccine; plant biocontrol agent.
XX
XX Acinetobacter baumannii.
XX
XX US6562958-B1.
XX
XX 13-MAY-2003.
XX
XX 04-JUN-1999; 99US-00328352.
XX
XX 09-JUN-1998; 98US-0088701P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Breton G, Bush D;
XX
XX WPI; 2003-576092/54.
XX
XX P-PSDB; ADA34150.
XX
XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
XX for diagnosing a bacterial disease, as components of antibacterial
XX vaccines, as targets for antibacterial drugs, or as biocontrol agents for
XX plants.
XX
XX Example; SEQ ID NO 1311; 328pp; English.
XX
XX The invention relates to isolated Acinetobacter baumannii nucleic acids.
XX The A. baumannii nucleic acids and polypeptides are useful as reagents
XX for diagnosing a bacterial disease, as components of antibacterial
XX vaccines, as targets for antibacterial drugs, to detect the presence of
XX A. baumannii and other Acinetobacter species in a sample, in screening
XX compounds for the ability to interfere with the A. baumannii life cycle
XX or to inhibit A. baumannii infection, and as biocontrol agents for
XX plants. The present sequence represents DNA encoding an A. baumannii
XX protein.
XX
XX Sequence 375 BP; 122 A; 64 C; 93 G; 96 T; 0 U; 0 Other;
XX
XX Query Match 10.5%; Score 85.8; DB 8; Length 375;
XX Best Local Similarity 57.4%; Pred. No. 6.5e-14;
XX Matches 175; Conservative 0; Mismatches 127; Indels 3; Gaps 1;
XX
XX 267 GCAATTGTCAGACCATGGAGATCCAGAGTTTCATCGGCTTTTACTGAAATCGGGATT 326
XX
XX 52 GCAATTGTAACCCGTTTAAATTTGGATGATGTGCGTGAAGCACTCTCTGACATTTGGTGA 111
XX
XX 327 CGAGGTGTTTACTGTTTCTCATGTGAGAGGTTTGGTGCACAGGAGGTTCTACCGAGAGA 386
XX
XX 112 CAAGGATTTACCGTAACTGAGTAAAGTTTGGTTCGTCACAAAGACATACAGAACTT 171
XX
XX 387 CACGGTGGCTCTGAGTCTCGGAAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTT 446
XX
XX 172 TACCGCGCGCTGAGTATGTTGTTGAT---TTCTTACCTAAAGTAAATAATCGAAATTCG 228
XX
XX 447 GTTAAGAAGACCAAGTGGAACTGTAATCAACACAATTAATTCGAGGAGCAAGGACAGGA 506
XX
XX -229 ATTAGTGATGAATGTCGACCGGCTAATTTGAGTCAATTACAGCTGTGGCAGCACTGGA 288
```

QY 507 GAGATTGGTGGGCAAGATTTTGGTTTCCCTGGTGTGAGATGTCATAGAGTTAGGACA 566
 DB 289 AAAATCGGCGACGGTAGATTTTGTGACCTAATCTGGACAGTCAATCCGTATCCGTACA 348
 QY 567 GGTGA 571
 DB 349 GGTGA 353

RESULT 8
 AAV21209_01
 Continuation (2 of 17) of AAV21209 from base 100001 (Methanococcus jannaschii circular
 WP Sequence split into 17 fragments LOCUS AAV21209 Accession Aav21209

Fragment Name	Begin	End
WP AAV21209_00	1	110000
WP AAV21209_01	100001	210000
WP AAV21209_02	200001	310000
WP AAV21209_03	300001	410000
WP AAV21209_04	400001	510000
WP AAV21209_05	500001	610000
WP AAV21209_06	600001	710000
WP AAV21209_07	700001	810000
WP AAV21209_08	800001	910000
WP AAV21209_09	900001	1010000
WP AAV21209_10	1000001	1110000
WP AAV21209_11	1100001	1210000
WP AAV21209_12	1200001	1310000
WP AAV21209_13	1300001	1410000
WP AAV21209_14	1400001	1510000
WP AAV21209_15	1500001	1610000
WP AAV21209_16	1600001	1664976

Query Match 10.4%; Score 84.6; DB 2; Length 110000;
 Best Local Similarity 56.0%; Pred. No. 1.5e-12;
 Matches 181; Conservative 0; Mismatches 139; Indels 3; Gaps 1;

QY 249 AAATTTTACAGGTGGGAAGCAATTTCTCAGACCATGGAGATCCAGCAAGTTTCATCGGCT 308
 DB 80891 AAAATGAAAAGTTTGAAGCAATCATAGACCGAGAGTTGGAGATTGTTAAAAGGCT 80950
 QY 309 TTAATGAAATCGGATTCGAGGTGTTTACTGTTTCTGATGTGAGAGGTTTGGTGCACAA 368
 DB 80951 TTGCTGATCTGATATGTTTGAATGACATGTTAGTGGGTTAAGGTTAGGGAGTTCAA 81010
 QY 369 GGAGGTCTACCGAGAGACACGGTGGCTCTGAGTCTCGGAGACAAATTTGTTGCTAAA 428
 DB 81011 GGTGGAAATGTTGAGAGGTATAGGGGAGAGATATTTGTTGAT---TTAATTCCAAAG 81067
 QY 429 GTTAAGATGAAATCGTTTGTAAAGAAAGACCAAGTGAATCTGTAATCAACACAAATATT 488
 DB 81068 GTTAAGATTGAATGTTTGTAAAGAGAGAGATGTTGATATGTTATATATATGC 81127
 QY 489 GAAGAGCAAGGACAGGAGAGATTTGGTATGGCAAGATTTTGGTCTGCTGTGAT 548
 DB 81128 GAGAAATGCAAGACAGCAAGCAAGCAAGGATGGAAAAATCTTCGTATACCAAGTAGAAGA 81187
 QY 549 GTCAATAGAGTTAGCAGAGTTGA 571
 DB 81188 GTCTAAGAGTAGAAGCAAAAGA 81210

RESULT 9
 AAV21209_14/c
 Continuation (15 of 17) of AAV21209 from base 1400001 (Methanococcus jannaschii circular
 WP Sequence split into 17 fragments LOCUS AAV21209 Accession Aav21209

Fragment Name	Begin	End
WP AAV21209_00	1	110000
WP AAV21209_01	100001	210000
WP AAV21209_02	200001	310000
WP AAV21209_03	300001	410000
WP AAV21209_04	400001	510000
WP AAV21209_05	500001	610000
WP AAV21209_06	600001	710000

Fragment Name	Begin	End
WP AAV21209_07	700001	810000
WP AAV21209_08	800001	910000
WP AAV21209_09	900001	1010000
WP AAV21209_10	1000001	1110000
WP AAV21209_11	1100001	1210000
WP AAV21209_12	1200001	1310000
WP AAV21209_13	1300001	1410000
WP AAV21209_14	1400001	1510000
WP AAV21209_15	1500001	1610000
WP AAV21209_16	1600001	1664976

Query Match 10.3%; Score 84.4; DB 2; Length 110000;
 Best Local Similarity 54.2%; Pred. No. 1.8e-12;
 Matches 194; Conservative 0; Mismatches 161; Indels 3; Gaps 1;

QY 230 TGATTTATTTCCAGACTCGAAATTTTACAAGGTGGGAAGCAATTTCTCAGACCATCGAAT 289
 DB 13404 TTATAAAATGAGTTCGAGATTATGAAAAGTTGAAGCAATCATAGACCGGAGAGTT 13345
 QY 290 CCAGCAAGTTTCTATCGGCTTTTACTGAAAAATCGGGAATTCGAGGTGTTACTGTTCTGATGT 349
 DB 13344 GGAGATTGTTAAAAGGCTTTTGTCTGATGCTGGGTATGTTGGAATGACTGTTAGTGAGGT 13285
 QY 350 GAGAGGTTTGGTGCACAGAGGTTCTACGAGACACGGTGGCTCTCGAGTTCTCGGA 409
 DB 13284 TAAGGGTAGGGAGTTTCAAGGTGGAATAGTTGAGAGGTATAGGGGGAGAGATATTTGT 13225
 QY 410 AGACAAATTTGTTCTAAAGTTAAGATGGAATCGTTTGTAAAGAAAGACCAAGTGGATC 469
 DB 13224 TGAT---TTAATTCCAAAGTTAAGATGAGTTGTTGTTAAAGAGGAGATGTTGATAA 13168
 QY 470 TGTAAATCAACACAAATAATTGAAGGAGCAAGAGAGAGATTTGGTGTATGCAAGATTTT 529
 DB 13167 TGTATTGATATCATATGCGAGATGCAAGAACAGAAACCCAGGAGATGCAAAAATCTT 13108
 QY 530 TGTTCCTGCTGTCAGATGTCATAGAGTTAGGACAGGTGAGCGTGGGAGAGAGCA 587
 DB 13107 CGTCATACCAAGTAGAAGAGTCTGTAAGAGTAAGAACAAAGAGAGAGAGAGAGCA 13050

RESULT 10
 ACF67367_52
 Continuation (53 of 57) of ACF67367 from base 5200001 (Photorhabdus luminescens nucleot.
 WP Sequence split into 57 fragments LOCUS ACF67367 Accession Aci67367

Fragment Name	Begin	End
WP ACF67367_00	1	110000
WP ACF67367_01	100001	210000
WP ACF67367_02	200001	310000
WP ACF67367_03	300001	410000
WP ACF67367_04	400001	510000
WP ACF67367_05	500001	610000
WP ACF67367_06	600001	710000
WP ACF67367_07	700001	810000
WP ACF67367_08	800001	910000
WP ACF67367_09	900001	1010000
WP ACF67367_10	1000001	1110000
WP ACF67367_11	1100001	1210000
WP ACF67367_12	1200001	1310000
WP ACF67367_13	1300001	1410000
WP ACF67367_14	1400001	1510000
WP ACF67367_15	1500001	1610000
WP ACF67367_16	1600001	1710000
WP ACF67367_17	1700001	1810000
WP ACF67367_18	1800001	1910000
WP ACF67367_19	1900001	2010000
WP ACF67367_20	2000001	2110000
WP ACF67367_21	2100001	2210000
WP ACF67367_22	2200001	2310000
WP ACF67367_23	2300001	2410000
WP ACF67367_24	2400001	2510000
WP ACF67367_25	2500001	2610000
WP ACF67367_26	2600001	2710000
WP ACF67367_27	2700001	2810000

WP ACF67367_28 2800001 2910000
 WP ACF67367_29 2900001 3010000
 WP ACF67367_30 3000001 3110000
 WP ACF67367_31 3100001 3210000
 WP ACF67367_32 3200001 3310000
 WP ACF67367_33 3300001 3410000
 WP ACF67367_34 3400001 3510000
 WP ACF67367_35 3500001 3610000
 WP ACF67367_36 3600001 3710000
 WP ACF67367_37 3700001 3810000
 WP ACF67367_38 3800001 3910000
 WP ACF67367_39 3900001 4010000
 WP ACF67367_40 4000001 4110000
 WP ACF67367_41 4100001 4210000
 WP ACF67367_42 4200001 4310000
 WP ACF67367_43 4300001 4410000
 WP ACF67367_44 4400001 4510000
 WP ACF67367_45 4500001 4610000
 WP ACF67367_46 4600001 4710000
 WP ACF67367_47 4700001 4810000
 WP ACF67367_48 4800001 4910000
 WP ACF67367_49 4900001 5010000
 WP ACF67367_50 5000001 5110000
 WP ACF67367_51 5100001 5210000
 WP ACF67367_52 5200001 5310000
 WP ACF67367_53 5300001 5410000
 WP ACF67367_54 5400001 5510000
 WP ACF67367_55 5500001 5610000
 WP ACF67367_56 5600001 5648894

Query Match 10.0%; Score 81.6; DB 7; Length 110000;
 Best Local Similarity 52.6%; Pred. No. 1.1e-11;
 Matches 202; Conservative 0; Mismatches 179; Indels 3; Gaps 1;

256 ACAAGTGGAGCAATTCGACACCATGGAGATCCAGCAAGTTTCATCGGCTTTACTGA 315
 49307 AAAGATTGATCGGATTATCAACCTTTCAATTAGATGATGCGTGAAGCTCTGGCG 49366
 316 AAATCGGATTGAGGTTACTGTTCTGATGTCAGAGGGTTTGGTGCACAAGGAGTT 375
 49367 AAGTGGGTATCACCGGAATCGGTGATGCAAGAGGTGAAGGTTTGGCGCCCAAAAGGTC 49426
 376 CTACCGAGAGCACGTCGCTGAGTTCTCGAGAGCAAAATTTGTTGCTAAAGTTAAGA 435
 49427 ATACAGAGCTGTATCGCGGTGAGAGATATATGTTGAT---TTTCTGCCAAAGTGAAA 49483
 436 TGGAAATCGTTTAAAGAACCAAGTGGATCTGTAATCAACACAATAATTGAAGGAG 495
 49484 TAGAAATTGTCGCCAGATGATTTGTCATACCTGTGTTGAACCATATTGCAAGCG 49543
 496 CAAGCAGAGGAGAGATTGGTATGCGCAAGATTTTTGGTCTGCTGATGATGTCATAA 555
 49544 CACAGACCGGAAATCGGTGATGTTAAATATTTGTTATTTGATGAGCAGCGTTTGTGC 49603
 556 GAGTTAGGACAGTCGCTGAGTTCTCGAGAGCAAGATGCTGCTGATGATGCTTTAC 615
 49604 GTATCCGACCGGTGAGCAGGATGAGAGCGGATTTAACTTAATTTTGGGCT 49663
 616 CGTCTTAGCAACAACAGAGCTCA 639
 49664 GGTGACAGCAGTAAATAAGCCA 49687

RESULT 11
 ACF65387_3
 Continuation (4 of 7) of ACF65387 from base 300001 (Photorhabdus luminescens nucleotide
 WP Sequence split into 7 fragments LOCUS ACF65387 Accession ACF65387
 WP Fragment Name Begin End
 WP ACF65387_0 1 110000
 WP ACF65387_1 100001 210000
 WP ACF65387_2 200001 310000
 WP ACF65387_3 300001 410000
 WP ACF65387_4 400001 510000

WP ACF65387_5 500001 610000
 WP ACF65387_6 600001 696798

Query Match 10.0%; Score 81.6; DB 7; Length 110000;
 Best Local Similarity 52.6%; Pred. No. 1.1e-11;
 Matches 202; Conservative 0; Mismatches 179; Indels 3; Gaps 1;

256 ACNAGTGGAGCAATTCGACACCATGGAGATCCAGCAAGTTTCATCGGCTTTACTGA 315
 38821 AAAGATTGATCGGATTATCAACCTTTCAATTAGATGATGCGTGAAGCTCTGGCG 88880
 316 AAATCGGATTGAGGTTACTGTTCTGATGTCAGAGGGTTTGGTGCACAAGGAGTT 375
 38881 AAGTGGGTATCACCGGAATGACAGTAACAGAGGTGAAGGTTTGGCGCCCAAAAGGTC 88940
 376 CTACCGAGAGCACGTCGCTGAGTTCTCGAGAGCAAAATTTGTTGCTAAAGTTAAGA 435
 88941 ATACAGAGCTGTATCGCGGTGAGAGATATATGTTGAT---TTTCTGCCAAAGTGAAA 88997
 436 TGGAAATCGTTTAAAGAACCAAGTGGATCTGTAATCAACACAATAATTGAAGGAG 495
 88998 TAGAAATTGTCGCCAGATGATTTGTCATACCTGTGTTGAACCATATTGCAAGCG 89057
 496 CAAGCAGAGGAGAGATTGGTATGCGCAAGATTTTTGGTCTGCTGATGATGTCATAA 555
 89058 CACAGACCGGAAATCGGTGATGTAATAATTTGTTATTTGATGAGCAGCGTTTGTGC 89117
 556 GAGTTAGGACAGTCGCTGAGTTCTCGAGAGCAAGATGCTGCTGATGATGCTTTAC 615
 89118 GTATCCGACCGGTGAGCAGGATGAGAGCGGATTTAACTTAATTTTGGGCT 89177
 616 CGTCTTAGCAACAACAGAGCTCA 639
 89178 GGTGACAGCAGTAAATAAGCCA 89201

RESULT 12
 ACF72027
 ID ACF72027 standard; DNA; 339 BP.
 XX
 AC ACF72027;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Photorhabdus luminescens nucleotide sequence #10494.
 XX
 KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 KW detection; food; gene expression; plant; animal; microorganism; toxin;
 KW antibiotic; biopesticide; virulence factor; disease model; plague;
 KW whooping cough; gene; ds.
 XX
 OS Photorhabdus luminescens.
 XX
 PN W0200294867-A2.
 XX
 PD 28-NOV-2002.
 XX
 PF 07-FEB-2002; 2002WO-IB0003040.
 XX
 PR 07-FEB-2001; 2001FR-00001659.
 XX
 PA (INSP) INST PASTEUR.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX
 XX Duchaud B, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
 PI Buchrieser C;
 XX
 XX WPI; 2003-148459/14.
 DR
 XX
 XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 XX useful e.g. as therapeutic antimicrobials and agricultural pesticides.
 XX
 XX Claim 2; SEQ ID NO 10494; 1205bp; French.

XX The invention relates to the isolation of genes and their encoded
 CC proteins from Photobacterium luminescens. The isolated sequences are
 CC sources of probes and primers for detecting the genome of P. luminescens
 CC and related species; to study polymorphisms; for gene analysis and for
 CC detection/amplification of the genes. Antibodies (Ab) raised against the
 CC polypeptides encoded by the genes are used for detection/identification
 CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 CC carry a gene-containing vector are used to select compounds that
 CC modulate, regulate, induce or inhibit expression of the genes in plants,
 CC animals or microorganisms other than P. luminescens and are able to alter
 CC response or sensitivity to toxins and antibiotics produced by P.
 CC luminescens. Cells transformed to express the genes are useful for
 CC recombinant production of the proteins, particularly toxins and
 CC antibacterial agents useful as insecticides, bactericides and fungicides. The
 CC genes, proteins, vectors containing the genes and Ab are also useful
 CC therapeutically (to treat microbial infection by bacteria or fungi that
 CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
 CC biopesticides. Other uses of the genes and the proteins are as virulence
 CC factors and for identifying targets of human diseases for which P.
 CC luminescens is a model (particularly plague and whooping cough). This
 CC sequence represents one of the isolated P. luminescens genes
 XX SQ Sequence 339 BP; 105 A; 51 C; 96 G; 87 T; 0 U; 0 Other;

Query Match 10.0%; Score 81.4; DB 7; Length 339;
 Best Local Similarity 55.0%; Pred. No. 1.1e-12;
 Matches 182; Conservative 0; Mismatches 146; Indels 3; Gaps 1;

QY 256 ACAAGTGGAGCAATGTCAGACCATGAGCAATCCAGCAAGTTTCATCGCTTACTGA 315
 DB 5 AAAAGATTGATGCGATTATCAACCTTCAAAATTAGATGATGCGTGAAGCTCTGCGG 64
 QY 316 AAATCGGATTCGAGGTGTACTCTCTGATGTCAGAGGTTTGGTGCAAGGAGGTT 375
 DB 65 AAGTGGTATCACCGGATGACATGACAGAGGTGAAAGTTTGGGCGCCAAAAGGTC 124
 QY 376 CTACCGAGAGACACGGTGGCTCTGAGTCTCGAAGCAAAATTTGTTGCTAAAGTTAGA 435
 DB 125 ATACAGAGCTGTATCGCGGTGTCAGAAATATATGTTGGAT---TTTCTGCCAAAAGTGA 181
 QY 436 TGGAAATCGTTTAAAGAGACCAAGTGGATCTGTAATCAACACATATTCAGGAG 495
 DB 182 TAGAAATGTCGTCGCCAGATGATGTCGATACCTGTGTTGAACCATTAATGACAGCG 241
 QY 496 CAACGACAGGAGATGTCGATGCGCAAGATTTTGTGTTTTCCTGCTGATGTCATGATA 555
 DB 242 CACAGACCGGAAATCGGTGATGTTAAATATTTGTTATTTGATGTCAGACGTTGTGTC 301
 QY 556 GAGTTAGACAGGTGAGCGTGGGAGAAAGC 586
 DB 302 GTATCCGACCGGTGAGCAGGATGAAGAGGC 332

RESULT 13
 AAT42063_03
 Continuation (4 of 19) of AAT42063 from base 300001 (Haemophilus influenzae complete genome)
 WP Sequence split into 19 fragments LOCUS AAT42063 Accession Aat42063

Fragment Name	Begin	End
WP AAT42063_00	1	110000
WP AAT42063_01	100001	210000
WP AAT42063_02	200001	310000
WP AAT42063_03	300001	410000
WP AAT42063_04	400001	510000
WP AAT42063_05	500001	610000
WP AAT42063_06	600001	710000
WP AAT42063_07	700001	810000
WP AAT42063_08	800001	910000
WP AAT42063_09	900001	1010000
WP AAT42063_10	1000001	1110000
WP AAT42063_11	1100001	1210000
WP AAT42063_12	1200001	1310000
WP AAT42063_13	1300001	1410000

WP AAT42063_14 1400001 1510000
 WP AAT42063_15 1500001 1610000
 WP AAT42063_16 1600001 1710000
 WP AAT42063_17 1700001 1810000
 WP AAT42063_18 1800001 1830121

Query Match 9.8%; Score 79.8; DB 2; Length 110000;
 Best Local Similarity 54.7%; Pred. No. 3.5e-11;
 Matches 181; Conservative 0; Mismatches 147; Indels 3; Gaps 1;

QY 256 ACAAGTGGAGCAATGTCAGACCATGAGCAATCCAGCAAGTTTCATCGCTTACTGA 315
 DB 63919 AAAAATCGAGCAATGATTAAACCTTTAAATTAGACGATGTCGAGAAAAGTCTTTTCAG 63978
 QY 316 AAATCGGATTCGAGGTGTACTCTCTGATGTCAGAGGTTTGGTGCAAGGAGGTT 375
 DB 63979 ATATTGTTATTCAGGTATGACATCAACAGAGTACCGGATTTGGTCTCAAAAGGTC 64038
 QY 376 CTACCGAGAGACACGGTGGCTCTGAGTCTCGAAGCAAAATTTGTTGCTAAAGTTAGA 435
 DB 64039 ATACAGACTTTATCGTGGTGGGCAATATATGTTGGAT---TTTCTGCCAAAAGTGA 64095
 QY 436 TGGAAATCGTTTAAAGAGACCAAGTGGATCTGTAATCAACACATATTCAGGAG 495
 DB 64096 TGGAAATGTTGTTCTCTGATGAGCTTGTGGATCAATGATTGAAGCGATTATTGAAAACGG 64155
 QY 496 CAACGACAGGAGATGTTGTTGTCGAGGATTTTGTGTTTTCCTGCTGTCAGATGTCATA 555
 DB 64156 CACAAACAGGTAAATCGTGGTGGGCAAAATTTTGTGTTTATCAGTTGAGAGCGCATCC 64215

QY 556 GAGTTAGACAGGTGAGCGTGGGAGAAAGC 586
 DB 64216 GCATTCGACCGGCGAGAAACAGGATGC 64246

RESULT 14

ABZ40022
 ID ABZ40022 standard; DNA; 363 BP.
 XX AC ABZ40022;
 XX AC ABZ40022;
 DT 07-MAR-2003 (first entry)
 XX N. gonorrhoeae nucleotide sequence SEQ ID 4633.
 DE N. gonorrhoeae nucleotide sequence SEQ ID 4633.
 KW Antibacterial; infection; vaccine; gene therapy; gene; ds.
 OS Neisseria gonorrhoeae.
 XX Neisseria gonorrhoeae.
 PN W0200279243-A2.
 XX W0200279243-A2.
 PD 10-OCT-2002.
 XX 10-OCT-2002.
 PF 12-FEB-2002; 2002WO-IB002069.
 XX 12-FEB-2002; 2002WO-IB002069.
 PR 12-FEB-2001; 2001GB-00003424.
 XX (CHIR-) CHIRON SPA.
 PA Fontana MR, Pizza M, Massignani V, Monaci E;
 PI WFI; 2003-058415/05.
 DR P-F5DB; ABP79052.
 XX PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
 PT medicament for treating or preventing N. gonorrhoeae infection.
 XX Disclosure; page 528; 815pp; English.

The present invention relates to proteins from Neisseria gonorrhoeae.
 CC Also disclosed are the nucleic acid molecules encoding the proteins and
 CC antibodies that specifically bind to the proteins. The composition
 CC comprising the protein, nucleic acid or antibody is useful for the

CC manufacture of a medicament for treating or preventing N. gonorrhoeae
CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records ABZ37706-ABZ42016 represent nucleic acid
CC molecules of the invention
XX
SQ Sequence 363 BP; 84 A; 77 C; 123 G; 79 T; 0 U; 0 Other;
Query Match 9.4%; Score 77; DB 7; Length 363;
Best Local Similarity 54.1%; Pred. No. 1.9e-11;
Matches 180; Conservative 0; Mismatches 150; Indels 3; Gaps 1;
QY 256 ACAAGTGGAGCAATGTTCAGACATCGAGCAATCCAGCAAGTTTCATCGGCTTTACTGA 315
Db 32 AAAAAATCGAGCGCATGTGTCAACCGTTCAAGCTCGACGACGTCCGCGAGCGGTTGACGG 91
QY 316 AAATCGGATTCGAGGTGTACTGTTCTGTGATGTGAGGGTTTGTGTCACAGAGAGTTT 375
Db 92 AAATCGGCAATTACGGGCATGACCGTCAAGAGGTCAAAGGGTTCCGCGAGGCGAAGGGGC 151
QY 376 CTACCGAGACACGGTGSCTCTGAGTCTCGGAGCAAAATTTGTTCTAAAGTTAAGA 435
Db 152 ATACGGAATCTATCGCGGTGCGAATACGCGGTGAT---TTCCTGCCAAGGTCAAAA 208
QY 436 TGGAAATCGTTTAAAGAAACACCAAGTGGATCTGTAATCAACACATAATTTGAAGGAG 495
Db 209 TCGATTTGTTGTGGCGGATGATCCGTGGACGGCGCATTTGACGTGATTTCGAGGTG 268
QY 496 CAAGCAGGAGAGATTGGTGATCGCAAGATTTTGTTCCTGCTGTGTCAGATGTCAATA 555
Db 269 CGCGTTCCGGGCAAAATCGCGACGCAAGATTTTGTGCTCGCGGTTCGAGGAGCAATCC 328
QY 556 GAGTTAGACAGGTGAGGTGGGAGAGAGCAG 588
Db 329 GTATCCGACGCGGCAACGTTTCGAGCGCGG 361
RESULT 15
AAA81472
ID AAA81472 standard; DNA; 16526 BP.
XX
AC AAA81472;
DT 04-DEC-2000 (first entry)
DE N. meningitidis partial DNA sequence gnm_20 SEQ ID NO:20.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB; ds.
XX
OS Neisseria meningitidis.
XX
FN WO20022430-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-US023573.
XX
PR 09-OCT-1998; 98US-0103794P.
PR 30-APR-1999; 99US-0132069P.
XX
PA (CHIR) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
XX
DR WPI: 2000-318079/27.
XX
PT Isolated nucleotide sequences of Neisseria meningitidis which can be used
PT in the diagnosis and treatment of N. meningitidis infection and other
PT Neisserial infections, for example, N.gonorrhoea.

Claim 7; Page 466-471; 1760pp; English.

PS The present invention describes methods of obtaining immunogenic proteins
XX from Neisseria genomic sequences. AAA81453 to AAA82414 represent
CC specifically claimed Neisseria meningitidis genomic DNA sequences;
CC AAA81260 to AAA81303 and AAB25620 to AAB25663 represent Neisseria DNA
CC sequences and their corresponding proteins; AAA81254 to AAA81259 and
CC AAA81304 to AAA81321 represent PCR primers used in the isolation of
CC Neisseria meningitidis DNA sequences; and AAA81322 to AAA81452 represent
CC Neisseria meningitidis MenB polynucleotide ORF sequences, which are all
CC used in the exemplification of the present invention. The nucleic acid
CC sequences, protein sequences, and antibodies against them, can be used in
CC the manufacture of a composition. The composition can be used as a
CC medicament (or in the manufacture of a medicament) for treating,
CC preventing or diagnosing infection due to Neisserial bacteria. For
CC example, some of the identified proteins could be components of vaccines
CC against Meningococcus B; against all serotypes; and/or against all
CC pathogenic Neisseriae. Identification of sequences from the bacterium
CC will also facilitate production of biological probes, particularly
CC organism-specific probes. Attempts to make efficacious Meningococcus B
CC vaccines have failed mainly due to antigen tolerance. Multivalent
CC vaccines have also been tried but none have successfully overcome
CC antigenic variability. The provision of further, complete sequences may
CC provide an opportunity to identify secreted or surface exposed proteins
CC that may be presumed targets for the immune system and which are not
CC antigenically variable or at least more conserved than other more
CC variable regions

XX SQ Sequence 16526 BP; 3183 A; 4006 C; 5067 G; 4268 T; 0 U; 2 Other;

Query Match 8.9%; Score 73; DB 3; Length 16526;
Best Local Similarity 55.3%; Pred. No. 1.3e-09;
Matches 184; Conservative 0; Mismatches 145; Indels 4; Gaps 2;

QY 256 ACAAGTGGAGCAATGTTCAGACATCGAGCAATCCAGCAAGTTTCATCGGCTTTACTGA 315
Db 13603 AAAAATCGAGCGCATTTGTCAAACCGTTCAAACCTCGACGCGGCGTTGACCG 13662
QY 316 AAATCGGATTCGAGGTGTACTGTTTCTGATGTGAGAGGGTTTGGTGCAAGAGGTT 375
Db 13663 AAATCGGATTCGAGGTGTACTGTTTCTGATGTGAGAGGGTTTGGTGCAAGAGG 13722
QY 376 CTACCGAGAGACACGGTGCCTCTGAGTTTTCGAGAGACAAATTTGTTGTAAGTTAAGA 435
Db 13723 ATACGGAATCTATCGCGCGCGGAATACGCGGTGAT---TTCCTGCCCAAAATCAAAA 13779
QY 436 TGGAAATCGTTTAAAGAAAGACCAAGTGGATCTGTATCAACACATAATTTGAGGAG 495
Db 13780 TCGAGCTGTTGTGGCGGATGATGCTGTGGAACGCGCGGATTCACGTGAT-TTCGAGGTG 13838
QY 496 CAAGGACAGGAGAGATTGGTGATGCGAAGATTTTGTGCTGTGTCAGATGTCAATA 555
Db 13839 CGCGTTCGGGAAATTCGCGACGCGCAGATTTTGTGCTGCGGTTGAGGAGGCAATCC 13898
QY 556 GAGTTAGACAGGTGAGCGTGGGAGAAAGCAG 588
Db 13899 GTATCCGACGCGGCAACGTTTCGAGCGCGGCGAG 13931

Search completed: May 27, 2004, 19:59:37
Job time : 289.934 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2004, 19:07:55 ; Search time 1992.7 Seconds
(without alignments)
12890.278 Million cell updates/sec

Title: US-09-756-541-13
Perfect score: 817
Sequence: 1 CTGAAAGTTGGTTAAAAA.....GTTCTTAATAAAAAAAA 817

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_ircv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rtd.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	499.6	61.2	566	9 AU236084	AU236084 AU236084
2	498.4	61.0	549	14 CB261070	CB261070 33-E9570-
3	459.6	56.3	685	14 CB261070	CD825321 BN25.060G
4	458.8	56.2	732	14 CB261070	CD820903 BN25.040D

5	456.8	55.9	745	14	CD819046
6	454	55.6	669	14	CD822497
7	369.6	45.2	451	9	AU226889
8	270.6	33.1	782	14	CB289081
9	262.8	32.2	723	13	BQ798001
10	252.8	30.9	732	9	AJ558383
11	249.6	30.6	812	14	CA922829
12	247.6	30.3	714	12	BI311072
13	246.6	30.2	597	10	BE239550
14	245	30.0	613	9	AW035791
15	245	30.0	702	12	BI932123
16	245	30.0	711	12	BI932336
17	245	30.0	736	12	BI932913
18	245	30.0	741	10	AW738071
19	245	30.0	804	12	BI930336
20	245	30.0	810	12	BI930060
21	244.8	30.0	612	10	BF645522
22	244	29.9	648	9	AI773079
23	243.4	29.8	630	10	AW929534
24	243.4	29.8	685	12	BI929601
25	241.8	28.6	799	12	BM113473
26	240.8	29.5	677	10	BE823431
27	240.2	29.4	756	12	BG595705
28	239.4	29.3	458	10	BE210153
29	239.4	29.3	458	10	BE210153
30	231.8	28.4	560	14	CA514765
31	231.8	28.4	579	10	AW223643
32	224.8	27.5	675	14	CB346213
33	224.6	27.5	866	13	CB1819
34	223.2	27.3	719	12	BJ573686
35	222.4	27.2	540	14	CB349912
36	221.8	27.1	551	14	CB349990
37	221.8	27.1	598	14	CB350294
38	221.6	27.1	657	14	CB349285
39	221.6	27.1	701	14	CB349370
40	219.2	26.8	540	10	AW624562
41	218.6	26.8	532	14	CB350222
42	213	26.1	564	13	BQ118918
43	213	26.1	729	14	CF475242
44	212.8	26.0	796	14	CF395143
45	211.8	25.9	673	14	CB085731

ALIGNMENTS

RESULT 1
AU236084
LOCUS AU236084 RAFL14 Arabidopsis thaliana cDNA clone RAFL14-63-K15 5', linear EST 01-APR-2002
DEFINITION AU236084 mRNA sequence.
ACCESSION AU236084
VERSION AU236084.1 GI:19875253
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 566)
AUTHORS Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M., Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y. and Shinozaki,K.
TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)
JOURNAL Unpublished (2002)
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp

CD819046 BN20.047L
CD822497 BN25.045G
AU226889 AU226889
CB289081 Y-B-112F0
BQ798001 EST 6939
AJ558383 AJ558383
CA922829 EST640547
BI311072 EST531282
BE239550 EST403599
AW035791 EST281945
BI932123 EST52012
BI932336 EST52225
BI932913 EST52802
AW738071 EST339498
BI930336 EST550225
BI930060 EST549949
BF645522 NF036F08E
AI773079 EST254179
AW929534 EST338322
BI929601 EST549490
BM113473 EST561009
BE823431 GM7000198
BG595705 EST494383
CF20467 gmtrRW3-
BE210153 SO39Q03.Y
CA514765 KS0903050
AW223643 EST300454
CB346213 CAB3SG000
CB1819 CB1819 C1C1
BJ573686 BJ573686
CB349912 CAB2SG000
CB349990 CAB2SG000
CB350294 CAB2SG000
CB349285 CAB2SG000
CB349370 CAB2SG000
AW624562 EST322507
CB350222 CAB2SG000
BQ118918 EST604494
CF475242 RTW2 14
CF395143 RTDS2_9_G
CB085731 bg15b09_9

An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pBluescript vector. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES

Location/Qualifiers
1..566
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL14-63-X15"
/tissue_type="root"
/lab_host="DH10B"
/clone_lib="RAFL14"
/note="Site_1: BamHI; Site_2: SalI"

ORIGIN

Query Match 61.2%; Score 499.6; DB 9; Length 566;
Best Local Similarity 98.2%; Pred. No. 4.8e-94;
Matches 536; Conservative 0; Mismatches 6; Indels 4; Gaps 3;
1 CTGAAGAGTTGTGTAAATAAACTAGATCATGGC-GGCGTCAATGACGAACCCATC 59
20 CTGAAGAGTTGTGTAAATAAACTAGATCATGGC-GGCGTCAATGACGAACCCATC 79
60 TCATTAACCTCTCTCGTTTCTATCTTCATGTCGAAGAACATGCTTCTCTGATTCAT 119
80 TCATTAACCTCTCTCGTTTCTATCTTCATGTCGAAGAACATGCTTCTCTGATTCAT 139
120 TCATTAACCTCTCTCGTTTCTATCTTCATGTCGAAGAACATGCTTCTCTGATTCAT 179
140 TCATTAACCTCTCTCGTTTCTATCTTCATGTCGAAGAACATGCTTCTCTGATTCAT 199
180 CCAGTAAATTAACAGTCTGTGTTTACCTGCTGTAGTCCCAATATCTTCTGATTCAT 239
200 CCAGTAAATTAACAGTCTGTGTTTACCTGCTGTAGTCCCAATATCTTCTGATTCAT 259
240 CCAGTAAATTAACAGTCTGTGTTTACCTGCTGTAGTCCCAATATCTTCTGATTCAT 299
260 CCAGTAAATTAACAGTCTGTGTTTACCTGCTGTAGTCCCAATATCTTCTGATTCAT 319
300 TCATCGGCTTTACTGAAATCGGANTCGAGGTGTACTGTTTCTGATTCGAGAGGTT 359
320 TCATCGGCTTTACTGAAATCGGANTCGAGGTGTACTGTTTCTGATTCGAGAGGTT 379
360 GGTGCAAGAGGTTTCTACCGAGGACACCGTCTGAGTTCTCGAAGACAAATTT 419
380 GGTGCAAGAGGTTTCTACCGAGGACACCGTCTGAGTTCTCGAAGACAAATTT 439
420 GTTCTCTAAAGTTAAGATGGAATCGTTTGTAAAGAACCAAGT-GGAATCTGTAATCAA 478
440 GTTCTCTAAAGTTAAGATGGAATCGTTTGTAAAGAACCAAGTGGGAATCTGTAATCAA 499
479 CACAATTAATGAAGAGCAAGG--ACAGGAGATGTTGATGCGCAAGATTTTCTTTG 536
500 CACAATTAATGAAGAGCAAGGACAGGAGAGATGTTGATGCGCAAGATTTTCTTTG 559
537 CCTGTG 542
560 CCTGTG 565

RESULT 2
CB261070
LOCUS
DEFINITION
33-E9570-012-004-A09-T7R MP12P-ADIS-012 Arabidopsis thaliana cDNA
clone MP12P769A0940 5-PRIME, mRNA sequence.
ACCESSION
CB261070
VERSION
CB261070.1 GI:32885843
KEYWORDS
EST.
SOURCE
Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 549)
Schmid,K.J., Soerensen,T.R., Stracke,R., Torjek,O., Altmann,T.,
Mitchell-Olds,T. and Weishaar,B.
Large-scale identification and analysis of genome-wide
single-nucleotide polymorphisms for mapping in Arabidopsis thaliana
Genome Res. 13 (6), 1250-1257 (2003)

TITLE

JOURNAL
MEDLINE
PUBMED

COMMENT

Contact: Weishaar B
ADIS DNA core facility at MPIZ
Max-Planck-Institute for Plant Breeding Research
Carl-von-Linne Weg 10, 50829 Koeln, Germany
Fax: 00492215062851
Email: weishaar@mpiz-koeln.mpg.de
Insert Length: 549 Std Error: 0.00
Plate: 4 row: A column: 09
Seq primer: T7R; CTAATACGACTCACTATAGGA.
Location/Qualifiers

FEATURES

source

1..549
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/cultivar="Landsberg erecta (Ler)"
/db_xref="GABI:589834"
/db_xref="taxon:3702"
/clone="MP12P769A0940"
/tissue_type="whole plant"
/dev_stage="adult plant, mixed stresses"
/lab_host="E. coli XL1-Blue MRF"
/clone_lib="MP12-ADIS-012"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; cDNA
library from Arabidopsis thaliana, accession Landsberg
erecta; six weeks old total plants grown under long-day
conditions in soil, whole adult plants were treated for 24
hours with different stresses, (1) at 4M-0 C in the dark,
(2), at 37 Grad C in the dark, (3) lying in the lab after
removing from soil, (4) in the greenhouse after wounding
leaves with a forceps, (5) in the lab watering with a 150
mM NaCl solution, (6) at 26 Grad C in the light/UV; equal
quantities of stressed plant material were pooled; library
was made at the Max-Planck-Institute for Plant Breeding
Research, Cologne, Germany; cloning sites SalI-NotI,
primer sites and orientation:
T7-SalI-CCACGCGTCG-5prime-cDNA-polyA-CC-NotI-SP6; Note:
Sequencing granted in the context of the GABI Arabidopsis
Verbund I: Genetic Diversity, 'Establishment of
high-efficiency SNP-based mapping tools and development of
methods for genome-wide mutation detection' PI: Bernd
Weishaar Sequence submission managed by RZPD/GABI-Primary
database: <http://gabi.rzpd.de>. This clone is available
from RZPD; contact RZPD (clone@rzpd.de) for further
information."

ORIGIN

Query Match 61.0%; Score 498.4; DB 14; Length 549;
Best Local Similarity 97.6%; Pred. No. 8.6e-94;
Matches 527; Conservative 0; Mismatches 11; Indels 2; Gaps 2;
14 TAAAAAAAACCTAGATCATGGCGGCTCAATGACGAACCCATCTCAATTAATCTCT 73
11 TAAAAAAAACCTAGATCATGGCGGCTCAATGACGAACCCATCTCAATTAATCTCT 70
74 CGTTTCTATTCTGATCGAAAGAACATGCTTCTCTGATTCGATTCGATTCGTCGG 133
71 CGTTTCTATTCTGATCGAAAG-AcATTGCTTCTCTGATTCGATTCGATTCGTCGG 129
134 ATTGAGATTCCTCCGACCATCTTGGCTCGATTTGGTTCACAAAGTACCAGTAATAACAG 193
130 ATTGAGATTCCTCCGACCATCTTGGCTCGATTTGGTTCACAAAGTACCAGTAATAACAG 189

```

QY 194 TCGTGTCTTACCTGTCGTAGTGGCCAAATATCTCTGATTATATATTCAGACTC-GAAT 252
DB 190 TCGTGTCTTACCTGTCGTAGAGCCAAAGCTCTCTGATTATATATTCAGACTCAGAAT 249
QY 253 TTACAAAGTGGAGCAATGTGACACCATCGAGAAATCCAGCAAGTTTTCATCGGCTTTAC 312
DB 250 TTACAAAGTGGAGCAATGTGACACCATCGAGAAATCCAGCAAGTTTTCATCGGCTTTAC 309
QY 313 TGAATAATCGGAATTCAGAGTGTACTGTTCTGATGTGAGAGGTTTGGTGCACAAGGAG 372
DB 310 TGAATAATCGGAATTCAGAGTGTACTGTTCTGATGTGAGAGGTTTGGTGCACAAGGAG 369
QY 373 GTTCTACCGAGACACCGTGGCTCTGAGTCTCGGAGACAAATTTGTTGCTAAAGTTA 432
DB 370 GTTCTACCGAGACACCGTGGCTCTGAGTCTCGGAGACAAATTTGTTGCTAAAGTTA 429
QY 433 AGATGGAATCGTCTTAAAGAAAGACCAAGTGGAAATCTGTAATCAACACAAATTTGAAG 492
DB 430 AGATGGAATCGTCTTAAAGAAAGACCAAGTGGAAATCTGTAATCAACACAAATTTGATG 489
QY 493 GAGCAAGACAGGAGAGATTTGGTGTGCAAGATTTTGTGTTGCTGTGTCAGATGCA 552
DB 490 GAGCAAGACAGGAGAGATTTGGTGTGCAAGATTTTGTGTTGCTGTGTCAGATGCA 549

RESULT 3
CD825321 LOCUS
DEFINITION BN25.060G1F011129 BN25 Brassica napus cDNA clone BN25060G17, mRNA
sequence.
ACCESSION CD825321
VERSION CD825321.1 GI:32507261
KEYWORDS EST.
SOURCE Brassica napus (rape)
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 685)
Genoplate.
Genoplate, a major partnership french program in plant genomics
Unpublished (2003)
Contact: Genoplate
Genoplate
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplate' (http://www.genoplate.com
and http://genoplate-info.infobiogen.fr).
FEATURES
Location/Qualifiers
1..685
/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="Jet neuf"
/db_xref="taxon:3708"
/clone="BN25060G17"
/tissue_type="seed"
/clone_lib="BN25"

ORIGIN
Query Match 56.3%; Score 459.6; DB 14; Length 685;
Best Local Similarity 85.9%; Pred. No. 9.5e-86;
Matches 534; Conservative 0; Mismatches 84; Indels 4; Gaps 2;

QY 29 AATCATGGCGGCTCAATGACGAACCCATCTCAATACTCTCTCGGTTCTATTCTGA 88
DB 5 AATCATGGCGGCTCAATGCGGAACCCATCTCTGCAACTCTCTCGGTTCTATTCTGA 64
QY 89 TCGAAG---AACAATGCTTTCTCTGATTGCAATTCGATTGTTGTTGGAATTCAGACATTC 145
DB 65 CCGAAGACACACATTCCTCTCTCAGATTGCATCTCGATATCTCTCGGTTCCGACATTC 124

```

```

QY 146 CGGACCATCTTGGCTCGATTGGTTCACAAAGTCACCGAGTAAATACAGTCGTGTTTACC 205
DB 125 CCACCGTCTTCCCTCGATTGGTTCGAAACCAACCGACCAAGGCTCGCGTTTCCCA 184
QY 206 TGTGTTAGTCCCAAAATATCTTCTGATTATATTCAGACTCGAAATTTTACAAGGTGGA 285
DB 185 TCTGTTAGAGCCCAAGCTCTTATGATTATATTCAGACTCAAAGTTTTTACAAGGTGGA 244
QY 286 AGCAATGTGACACCATGAGAAATCCAGCAAGTTTTCATCGGCTTTTACTGAAATCGGAT 325
DB 245 AGCAATGTGACAGCCATGAGAAATCCAGCAATGTTTCATCGGCTTTTACTGAAATCGGAT 304
QY 326 TCGAGTGTGTTACTGTTCTCTGATGTGAGAGGTTTGGTGCACAAGGAGTTTCTACCGAGAG 385
DB 305 TCGAGTGTGACGHTTCTGATGTGAGAGGTTTGGTGCACAAGGTTGATCCAAAGAGAG 364
QY 386 ACACGCTGCTCTGAGTTCTCGAAGACAAATTTGTTGCTAAAGTTTAAGATGGAATCGT 445
DB 365 ACATGCTGCTCTGAGTTCTCTGAAAGACAAGTTGCTTAAAGTTTAAGATGGAATCGT 424
QY 446 TTTTAAAGAAAGACCAAGTGGAAATCTGTAATCAACACAAATTAAGAGGACGAGGACAGG 505
DB 425 TTTTAAAGAAAGACCAAGTGGAGTCTGTAATCAACACAAATTAAGATGGAATCGATGGAAGAACAGG 484
QY 506 AGAGATTCGTGATGCAAGATTTTGTGTTGCTGTGTCAGATGTCATAAGAGTTAGGAC 565
DB 485 AGAGATTCGAGATGCGAATTAATTTGTTTGTGCTGTGTCAGATGTCATAAGAGTTAGGAC 544
QY 566 AGGTGAGCTGCGGAGAAAGACAGAGAGATGATCGTGATATGCTTTTACCGCTTTAGGA 625
DB 545 AGGTGAGCTGCGGAGAAAGACAGAGAGATGATCGTGATATGCTTTTACCGCTTTAGGC 604
QY 626 ACAAAC-AGAGCTCAAGATGG 646
DB 605 ACAAACACAGAGCTTAAGATGG 626

RESULT 4
CD820903 LOCUS
DEFINITION BN25.040D05F011026 BN25 Brassica napus cDNA clone BN25040D05, mRNA
sequence.
ACCESSION CD820903
VERSION CD820903.1 GI:32502843
KEYWORDS EST.
SOURCE Brassica napus (rape)
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 732)
Genoplate.
Genoplate, a major partnership french program in plant genomics
Unpublished (2003)
Contact: Genoplate
Genoplate
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplate' (http://www.genoplate.com
and http://genoplate-info.infobiogen.fr).
FEATURES
Location/Qualifiers
1..732
/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="Jet neuf"
/db_xref="taxon:3708"
/clone="BN25040D05"
/tissue_type="seed"
/clone_lib="BN25"

ORIGIN
source

```

Query Match 56.2%; Score 458.8; DB 14; Length 732;
 Best Local Similarity 80.8%; Pred. No. 1.4e-85;
 Matches 609; Conservative 0; Mismatches 112; Indels 33; Gaps 5;

QY 29 AATCATGCGCGGCAATGAGCAAGCAACCCATCTCAATAACTTCTCTCGGTTTCTAATTCGA 88
 Db 8 AATCATGCGCGGTTCAATCGGGAACCCATCTCTGCAACCTCTCTCGGTTTCTAATTCGA 67

QY 89 TCGAAG---AACATTCCTTCTCTGATTCGATTCGATTCGATTCGATTCGATTCGATTC 145
 Db 68 CCGAAGAAACAACATTCCTCTCAGATTCGATTCGATTCGATTCGATTCGATTCGATTC 127

QY 146 CCGACCAATCTTGCCTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 205
 Db 128 CCGCCCGCTCTTCCCTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 187

QY 206 TGTGCTTAGTGCCTCAATATCTCTGATTCGATTCGATTCGATTCGATTCGATTCGATTC 265
 Db 188 TCTGCTTAGAGCCCAAGCTCTTATGATTCGATTCGATTCGATTCGATTCGATTCGATTC 247

QY 266 AGCAATTCGACGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 325
 Db 248 AGCAATTCGACGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 307

QY 326 TCGAGGTGTTACTTCTCTGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 385
 Db 308 TCGAGGTGTTACTTCTCTGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 367

QY 386 ACAGGTGGCTCTGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 445
 Db 368 ACATGTTGGCTCTGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 427

QY 446 TGTAAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 505
 Db 428 TGTAAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 487

QY 506 AGATGTTGGTGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 565
 Db 488 AGATGTTGGTGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 547

QY 566 AGGTGAGCGTGGGAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 625
 Db 548 AGGTGAGCGTGGGAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 607

QY 626 ACAAAC-AGAGCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 684
 Db 608 ACAAACAGAGCTTAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 640

QY 685 TAAATCAATGAGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 744
 Db 641 AATGAATGAGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 699

QY 745 TGTACATGAATATGCGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 778
 Db 700 TGTACATGAATGATGCA-ACAATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 732

RESULT 5
 CD819046
 LOCUS
 DEFINITION BN20.047L06F01.1226 BN20 Brassica napus cDNA clone BN20047L06, mRNA
 sequence.
 CD819046
 VERSION CD819046.1 GI:32500986
 KEYWORDS EST.
 SOURCE Brassica napus (rape)
 ORGANISM Brassica napus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 745)
 Genoplante.
 Genoplante, a major partnership french program in plant genomics

Unpublished (2003)
 Contact: Genoplante
 Genoplante
 93, rue Henri Rochefort 91025 EVRY CEDEX France
 Tel: 33 1 69 47 54 00
 Fax: 33 1 69 47 54 10
 This sequence has been generated in the framework of the french
 plant genomics programme 'Genoplante' (<http://www.genoplante.com>)
 and <http://genoplante-info.infobiogen.fr>.

FEATURES
 Location/Qualifiers
 1..745
 /organism="Brassica napus"
 /mol_type="mRNA"
 /cultivar="Jet neuf"
 /db_xref="taxon:3708"
 /clone="BN20047L06"
 /tissue_type="seed"
 /clone_id="BN20"

ORIGIN

Query Match 55.9%; Score 456.8; DB 14; Length 745;
 Best Local Similarity 82.0%; Pred. No. 3.5e-85;
 Matches 551; Conservative 0; Mismatches 117; Indels 4; Gaps 2;

QY 29 AATCATGCGCGGCAATGAGCAAGCAACCCATCTCAATAACTTCTCTCGGTTTCTAATTCGA 88
 Db 68 AATCATGCGCGGTTCAATCGGGAACCCATCTCTGCAACCTCTCTCGGTTTCTAATTCGA 127

QY 89 TCGAAG---AACATTCGCTTCTCTGATTCGATTCGATTCGATTCGATTCGATTCGATTC 145
 Db 128 CCGAAGAAACAACATTCCTTCTCAGATTCGATTCGATTCGATTCGATTCGATTCGATTC 187

QY 146 CCGACCAATCTTGCCTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 205
 Db 188 CCGCCCGCTTCTCCCTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 247

QY 206 TGTGCTTAGTGCCTCAATATCTCTGATTCGATTCGATTCGATTCGATTCGATTCGATTC 265
 Db 248 TCTGCTTAGAGCCCAAGCTCTTATGATTCGATTCGATTCGATTCGATTCGATTCGATTC 307

QY 266 AGCAATTCGACGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 325
 Db 308 AGCAATTCGACGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 367

QY 326 TCGAGGTGTTACTTCTCTGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 385
 Db 368 TCGAGGTGTTACTTCTCTGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 427

QY 386 ACAGGTGGCTCTGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 445
 Db 428 ACATGTTGGCTCTGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 487

QY 446 TGTAAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 505
 Db 488 TGTAAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 547

QY 506 AGATGTTGGTGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 565
 Db 548 AGATGTTGGTGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 607

QY 566 AGGTGAGCGTGGGAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 625
 Db 608 AGGTGAGCGTGGGAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 667

QY 626 ACAAAC-AGAGCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 684
 Db 668 ACAAACAGAGCTTAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 727

QY 685 TAAATCAATG 696
 Db 728 TGGTCTGAGTG 739

```

RESULT 6
CDB22497
LOCUS      CD822497               669 bp    mRNA    linear    EST 10-JUL-2003
DEFINITION BN25.045G20F0201C8 BN25 Brassica napus cDNA clone BN25045G20, mRNA
sequence.
ACCESSION  CD822497
VERSION     CD822497.1
KEYWORDS   EST.
SOURCE      Brassica napus (rape)
ORGANISM   Brassica napus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1 (bases 1 to 669)
AUTHORS    Genoplante, a major partnership french program in plant genomics
            Unpublished (2003)
TITLE      Genoplante, a major partnership french program in plant genomics
JOURNAL    Unpublished (2003)
COMMENT    Contact: Genoplante
            Genoplante
            93, rue Henri Rochefort 91025 EVRY CEDEX France
            Tel: 33 1 69 47 54 00
            Fax: 33 1 69 47 54 10
            This sequence has been generated in the framework of the french
            plant genomics programme "Genoplante" (http://www.genoplante.com)
            and http://genoplante-info.inbio.gen.fr.
            Location/Qualifiers
            1..669
            /organism="Brassica napus"
            /mol_type="mRNA"
            /cultivar="Jef neuf"
            /db_xref="taxon:3708"
            /clone="BN25045G20"
            /tissue_type="seed"
            /clone_lib="BN25"

FEATURES             source
    source
    Query Match      55.6%; Score 454; DB 14; Length 669;
    Best Local Similarity 85.6%; Pred. No. 1.4e-84;
    Matches 529; Conservative 0; Mismatches 85; Indels 4; Gaps 2;

QY 33 ATGCGCGGCTCAATGAGCAAGCAACCCCTCAATCACTTCTCGGTTTCTATTCTGAPCGA 92
DB 2 ATGCGCGGCTCAATGAGCAAGCAACCCCTCAATCACTTCTCGGTTTCTATTCTGACCGCA 61
QY 93 AAG---AACATTCCTTCTCTGATTTGCAATTCGATTTGTTCTGATTTCAAGATTCACATTCGCA 149
DB 62 AAGAACAACATTCCTTCTGATTTGCAATTCGATTTGTTCTGATTTCAAGATTCACATTCGCA 121
QY 150 CCAATTTGCTCGATTTGTTGTCACAAAGTCACCGAGTAATACAGTCGTGTTTACCTGTC 209
DB 122 CGCTCTTCCTCGATTTGTTGTCACAAAGTCACCGAGTAATACAGTCGTGTTTACCTGTC 181
QY 210 GTTAGTGCCCAATATCTTCTGATTTATATTCAGACTCGAATTTTACAGGTGGAAGCA 269
DB 182 GTTAGAGCCCAAGCTCTTATGATTAATATTCAGACTCGAATTTTACAGGTGGAAGCA 241
QY 270 ATTGTGACACCAATGGAGATTCAGCAAGTTTCAATCGCTTACTGAAATCGGATTCGA 329
DB 242 ATTGTGACGCAATGGAGATTCAGCAAGTTTCAATCGCTTACTGAAATCGGATTCGA 301
QY 330 GGTGTATCTGTTTCTGATGAGAGGGTTTGGTGCAACAGGAGTTCTACCGAGAGACAC 389
DB 302 GGTGTGACGGTTTCTGATGTCGAGGGTTTGGTGCAACAGGAGTTCTACCGAGAGACAC 361
QY 390 GGTGCTCTGAGTTCTCGAAGCAAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTT 449
DB 362 GGTGGTCTGAGTTCTCTGAAGCAAGTTCTGTTCTAAAGTTAAGATGAGATCGTTGTT 421
QY 450 AAGAAGACCAAGTGAATCTGTAATCAACAAATTTGAAGGACCAAGGACAGAGAG 509
DB 422 AAGAAGACCAAGTGAATCTGTAATCAACAAATTTGAAGGACCAAGGACAGAGAG 481
QY 510 ATTGGTGATGGCAAGATTTTGTGCTGTGTCAGATGTCATTAAGAGTTAGGACAGGT 569

```

```

Db 482 ATCGAGATGGCAAAATATTCGTTTTCCTGCTGTGTCAGATGTCATAAGGGTTCGACAGGT 541
QY 570 GAGCGTGGGGAGAGAGCAGAGAGATGACTGGTGATATGCTTTCACCGCTTAGGACAA 629
DB 542 GAGCGTGGAGACAGCAGAGAGATGACTGGTGATATGCTTTCGTCATCTTAGGACAA 601
QY 630 AC-AGAGCTCAAGATGG 646
DB 602 ACAAGAGCTTAAGATGG 619

RESULT 7
AU226889/c
LOCUS      AU226889             451 bp    mRNA    linear    EST 23-APR-2002
DEFINITION AU226889 RAFL14 Arabidopsis thaliana cDNA clone RAFL14-63-K15 3',
            mRNA sequence.
ACCESSION  AU226889
VERSION     AU226889.1
KEYWORDS   EST.
SOURCE      Arabidopsis thaliana (thale cress)
ORGANISM   Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1 (bases 1 to 451)
AUTHORS    Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
            Akiyama,K., Enju,A., Oono,Y., Sakurai,T., Carninci,P., Kawai,J.,
            Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A.,
            Muramatsu,M., Hayashizaki,Y. and Shinozaki,K.
            Large scale analysis of Arabidopsis full-length cDNA
            Unpublished (2002)
            Contact: Motoaki Seki
            Plant Functional Genomics Research Group
            RIKEN Genomic Sciences Center
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: 81-298-36-4359
            Fax: 81-298-36-9060
            Email: msekic@tc.riken.go.jp
            An Arabidopsis full-length cDNA library was constructed essentially
            as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
            and XhoI was ligated to modified lambda f1c-1 vector (Carninci et
            al., submitted for publication) digested with BamHI and SalI. This
            clone is in a modified pBluescript vector. Please visit our web
            site (http://www.gsc.riken.go.jp/e/plant/index\_e.html) for further
            details.
            Location/Qualifiers
            1..451
            /organism="Arabidopsis thaliana"
            /mol_type="mRNA"
            /db_xref="taxon:3702"
            /clone="RAFL14-63-K15"
            /tissue_type="root"
            /lab_host="DH10B"
            /clone_lib="RAFL14"
            /note="Site_1: BamHI; Site_2: SalI"

FEATURES             source
    source
    Query Match      45.2%; Score 369.6; DB 9; Length 451;
    Best Local Similarity 96.9%; Pred. No. 6.3e-67;
    Matches 409; Conservative 0; Mismatches 9; Indels 4; Gaps 3;

QY 391 GTGGCTCTGAGTTCTCGAAGCAAAATTTGTTCTAAAGTTAAGATGGAATCGTTGTTA 450
DB 450 GTGGCTCTGAGTTCTC-GAAGCAAAATTTGTTCTAAAGTTAAGATGGAATCGTTGTTA 392
QY 451 AGAAGACCAAGTGGAAATCTGTAATCAACAAATTTGAAGGACCAAGGACAGAGAGA 510
DB 391 AGAAGACCAAGTGGAAATCTGTAATCAACAAATTTGAAGGACCAAGGACAGAGAGA 333
QY 511 TTGGTGATGGCAGATTTTGTGCTGTGTCAGATGTCATAAGAGTTAGGACAGGTG 570
DB 332 TTGGTGATGGCAAAATTTTGTGCTGTGTCAGATGTCATAAGAGTTAGGACAGGTG 273

```

QY 571 AG--CGTGGGAGAAAGCAGAGAGATGACTGGTGTATGCTTTCACCGCTTAGGACGA 628
 |||||
 Db 272 AGCGGTGGGGGAGAAAGCAGAGAGATGCTGGTGTATGCTTTCACCGCTTAGGACGA 213
 |||||
 QY 629 ACAGAGCTCAAGATGGT 688
 |||||
 Db 212 ACAGAGCTCAAGATGGT 153
 |||||
 QY 689 AATGAATGAGT 748
 |||||
 Db 152 AATGAATGAGT 93
 |||||
 QY 749 CATGAATATGAGAAACATCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 808
 |||||
 Db 92 CATGAATATGAGAAACATCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 33
 |||||
 QY 809 AA 810
 |||||
 Db 32 AA 31

RESULT 8
 CB289081
 LOCUS
 DEFINITION V-B-1:2F06 VAN-Baker-1 Vitis aestivalis cdNA clone V-B-112F06 5', mRNA sequence.

ACCESSION CB289081.1 GI:28602822

VERSION EST.

KEYWORDS

SOURCE

ORGANISM

Vitis aestivalis

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; Vitaceae; Vitis.

1 (Bases 1 to 782)

Hou, H.S., Phanikanth, T.V., Kovacs, L. and Qiu, W.P.

Expressed sequence tags of young leaf tissues of a

disease-resistant Vitis aestivalis var. Norton

Unpublished (2003)

Contact: Wenping Qiu

Department of Fruit Science

Southwest Missouri State University-Mountain Grove

9740 Red Spring Road, Mountain Grove, MO 65711-2999, USA

Tel: 417 926 4105

Fax: 417 926 6646

Email: wq070f@msu.edu

Insert Length: 782 Std Error: 0.00

Plate: VAN-Baker-1-12 row: F column: C6

Seq primer: T3 PRIMER

High quality sequence stop: 782

POLYA-No.

Location/Qualifiers

1. 782

/organism="Vitis aestivalis"

/mol type="mRNA"

/cultivar="Norton"

/db xref="taxon:3605"

/clone="V-B-112F06"

/tissue_type="Leaf"

/dev stage="Young leaf"

/lab host="XL10-Gold E.coli"

/clone lib="VAN-Baker-1"

/notes="Vector: pBluescript II SK (+) Phagemid; Site 1: Xho

I; Site 2: EcoR I; VAN-Baker-1 is a cDNA library of Norton

grape young leaves (Vitis. aestivalis var. Norton). Norton

grapevines were grown under normal greenhouse conditions.

The cDNA synthesis and library construction was performed

according to the instruction manual for pBluescript II XR

cDNA library construction kit provided by Stratagene."

ORIGIN

source

33.1%; Score 270.6; DB 14; Length 782;

Query Match

Best Local Similarity 72.7%; Pred. No. 2.1e-46;
 Matches 364; Conservative 0; Mismatches 134; Indels 3; Gaps 1;
 QY 103 CTTTCTCTGATTGCAATTCGATTGTTCTGGATTTCAGACATTCGACCATCTTCGCTCG 162
 |||||
 Db 74 CTCTCATAGATGACATCTTGTGATCCGCCCAATCTCAAGATTCTCGAAATTTCCAGTTTA 133
 |||||
 QY 163 ATTTGTGTCACAAAGTCCCGAGTAAATACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 222
 |||||
 Db 134 ATTTGTGCTTAA--CCGCTCAAAGATGATCTGTCTCTCCATTTGTCAGAGCTCAGA 190
 |||||
 QY 223 TATCTCTGATTATATTCAGACTCGAAATTTTCAAGGTGGAGCAATTTGTCCAGCCAT 282
 |||||
 Db 191 GCTCTCCAGATTATACCCAGACGCGCAGTTTACAAAGTTGAAAGCATCTCAGGCCCT 250
 |||||
 QY 283 GGAGAAATCCAGCAAGTTTTCATCGGCTTTTACTGAAATTCGGATTTCGAGGTGTGTGTGTGT 342
 |||||
 Db 251 GCGGAATCCAGCAGGTTTCTCGGCTTTCTGAAATGGGTATTCGTGGTGTGTGTGTGTGT 310
 |||||
 QY 343 CTGATGTGAGAGGGTTTGTGTGCAAGAGGTTTCTACCGAGAGACACGGTGGCTCTGAGT 402
 |||||
 Db 311 CTGATGTCCAGGCTTTGGTGTCTCAAGCGGATCGCTGAAGACAGGCGAGCTCTGAAT 370
 |||||
 QY 403 TCTCGGAGACAAATTTGTTGCTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAG 462
 |||||
 Db 371 TTTCTGAGGACAAATTTGTTGCTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAGTTTAAAG 430
 |||||
 QY 463 TGGAAATCTGTAATCAACAAATTTGAGGAGCAGGACGAGAGAGATGTTGTTGATGGCA 522
 |||||
 Db 431 TTGAGCGATATTTCAGAGATCAATGAGTGGCAGGACTGGAGAGATTGGTATGGCA 490
 |||||
 QY 523 AGATTTTGTGCTGTGTGATGTCATGATGTCATGATGTCATGATGTCATGATGTCATGATGTCAT 582
 |||||
 Db 491 AGATCTTTTGTGCTGTCATGATGTCATGATGTCATGATGTCATGATGTCATGATGTCATGATGTCAT 550
 |||||
 QY 583 AAGCAGACAGATGCTGCTGTG 603
 |||||
 Db 551 TGGCTGAGAGGATGACTGCTGTG 571

RESULT 9

BO798001/c

LOCUS

DEFINITION EST 6939 Ripening Grape berries lambda Zap II Library Vitis

vinifera cdNA clone RT93C02 3', mRNA sequence.

ACCESSION BO798001

VERSION BO798001.1 GI:22012967

KEYWORDS EST.

SOURCE

ORGANISM

Vitis vinifera

Vitis vinifera

REFERENCE

AUTHORS

Abbal, P., Agasse, A., Ageorges, A., Atanassova, R., Barrieu, F.,

Couture, C., Dedaldechamp, F., Delrot, S., Glissant, D., Grimplet, J.,

Hamdi, S., Romieu, C. and Terrier, N.

Generation of Expressed Sequence Tag from Grape Berry (skin, pulp

or seeds) at Various Developmental Stages

Unpublished (2002)

Contact: Romieu C.

Unite de Recherche des Produits de la Vigne

Institut National de la Recherche Agronomique

2, place Viala, 34 060 Montpellier Cedex 01, France

Tel: 00-33-(0)4-99-61-28-62

Fax: 00-33-(0)4-99-61-28-57

Email: romieu@ensan.inra.fr

Seq primer: T7.

Location/Qualifiers

1. 723

/organism="Vitis vinifera"

/mol type="mRNA"

/cultivar="Shiraz"

```
/db_xref="taxon:29760"
/clone="RT093C02"
/dev_stage="ripening stage"
/clone_lib="Ripening Grape berries Lambda Zap II Library"
/note="Organ: Fruit; Vector: Lambda Zap II; Site 1: Eco
RI; Site 2: XhoI; Oriented library, construction described
in Generation of ESTs from grape berry (skin, pulp or
seeds) at various developmental stages by Terrier, N.
Ageorges, A., Abbal, P., Romieu, C. in J. Plant Physiol. 158
(12): 1575-83 2001"

ORIGIN
Query Match 32.2%; Score 262.8; DB 13; Length 723;
Best Local Similarity 76.8%; Pred. No. 9.3e-45;
Matches 321; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

Qy 186 AATAACAGTCGTGTTTACCTGTGCTGTTAGTGCCTCAATATCTCTGATTATATCCAGACTCGAAATTTTACAAAGTG 263
Db 675 AAGAATGGATCTGTTCTTCCCATGTCAGAGCTCAGAGCTCTCCAGATTATACCCAGAC 616

Qy 246 TCGAAATTTACAGGTGGAAGCAATTCAGACCATGAGAAATCCAGAGTTTCATCG 305
Db 615 GCGCAGTTTTCACAAAGTTGAACGATCTGAGGCCCTCGCAATCCAGCAGGTTTCCTCG 556

Qy 306 GCCTTACTGAAATCGGATTCGAGGTGCTTACTGTTCTGATGTCAGAGGGTTTCGTGCA 365
Db 555 GCCTTCTGAAATCGGATTCGAGGTGCTTACTGTTCTGATGTCAGAGGGTTTCGTGCT 496

Qy 366 CAAGAGAGTTCTACCGAGAGACACGCTGCTCTGAGTTCTCGAAGACAAATTTGTTGCT 425
Db 495 CAAGCGGTTTCGCTGAAAGACAGCAGCGCTCTGAAATTTCTGAGACAAATTTGTTGCT 436

Qy 426 AAGTTAGATCGAAATCGTTTGAAGAAAGACCAAGTGAATCTGTTATCAACAAATA 485
Db 435 AAGTCAAAATGGAATGTTGCTGAGCAAGACCAAGTTGAAGCAAGTAAATGCAAGATC 376

Qy 486 ATTCAAGAGCAAGACAGGAGAGATTCGTGATGCAAGATTTTGTGCTGCTGCA 545
Db 375 AATGAGTGGCAAGACCTGGAGAGATTCGTGATGCAAGATTTTGTGCTGCTGCAATCG 316

Qy 546 GATCTCATAGAGTTAGACAGAGGTGAGCGTGGGGGAAAGACAGAGATGACCTGGTG 603
Db 315 GATGTGATAGAGTTCCGACCGGTGACGCTGAGAGCTGGCTGAGAGATGACCGGGTG 258

RESULT 10
AJ558383
LOCUS AJ558383 Antirrhinum majus whole plant Antirrhinum majus cDNA clone
DEFINITION 018_1_07_f22, mRNA sequence.
ACCESSION AJ558383.1 GI:31660955
VERSION AJ558383.1
KEYWORDS EST.
SOURCE Antirrhinum majus (snapdragon)
ORGANISM Antirrhinum majus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Antirrhinaceae; Antirrhineae;
Antirrhinum.
1 (bases 1 to 732)
Zachgo, S., Stueber, K., Saedler, H., Sommer, H. and Schwarz-Sommer, Z.
Antirrhinum EST collection
Contact: Schwarz-Sommer Z
Molekulare Pflanzen-genetik
MPI fuer Zuechtungs-forschung
Carl-von-Linne Weg 10, D-50829, Germany.
Location/Qualifiers
1..732
/organism="Antirrhinum majus"
/mol_type="mRNA"
/db_xref="taxon:4151"
/clone="018_1_07_f22"

/db_xref="taxon:29760"
/clone="RT093C02"
/dev_stage="ripening stage"
/clone_lib="Ripening Grape berries Lambda Zap II Library"
/note="Organ: Fruit; Vector: Lambda Zap II; Site 1: Eco
RI; Site 2: XhoI; Oriented library, construction described
in Generation of ESTs from grape berry (skin, pulp or
seeds) at various developmental stages by Terrier, N.
Ageorges, A., Abbal, P., Romieu, C. in J. Plant Physiol. 158
(12): 1575-83 2001"

ORIGIN
Query Match 30.9%; Score 252.8; DB 9; Length 732;
Best Local Similarity 77.0%; Pred. No. 1.1e-42;
Matches 308; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

Qy 204 CCTGCTGCTAGTGGCCCAATATCTCTGATTATATATCCAGACTCGAAATTTTACAAAGTG 263
Db 172 CCAATGTTAGGGCTCAGATTTCTCCAGATCCCTACCTGACTCCAAAGTTCTACAAGTC 231

Qy 264 GAAGCAATTTGTCAGACCATGAGAAATCCAGCAAGTTTTCATCGCTTTACTGAAAAATCGGG 323
Db 232 GAAGCTATTTAAGACCTCGAGGATCCACAGAGTTTCTCTCGGCTTCTGAAAAATGGGA 291

Qy 324 ATTGAGGTGTTACTGTTTCTGATGTCAGAGGGTTTGGTGGCAAGAGGTTCTACCGAG 383
Db 292 ATTGCTGGTGTCTCTGATGTCAGAGGGTTTGGTGTCTCAGGGTGTGTTGACAGAA 351

Qy 384 AGACAGCGTGGCTGCTGAGTTCTCGGAGAGCAAAATTTGTTGCTAAAGTTAAGATGGAATC 443
Db 352 AGACAGTTTGGCTCTGAAATTTCTCAGAAAGCAAAATTTGTTGCAAAAGTTAAGATGGAAT 411

Qy 444 GTTGTTAAGAAAGACCAAGTGGAAATCTGTAATCAACACAAATATTCAGAGCAAGGACA 503
Db 412 GTTGAGCAAGACACAGGTTGATGGGTAAAGAGATTAATAGAGCAAGGACAAGACA 471

Qy 504 GGAGAGATTGGTGTAGTGGCAAGATTTTGTGTTTGGCTGCTCAGATGTCATAAGATTAGG 563
Db 472 GGAGAGATTGGTGTAGTGGCAAGATTTTGTGCTTCCAGTTTCTGATGTAATAGAGTTGCG 531

Qy 564 ACAGTGTGAGCTGGCGAGAAAGCAGAGAGATGACTGGTG 603
Db 532 ACTGCTGAACGAGGGGAAAGGCTGAGAGGATGGTGGAG 571

RESULT 11
CA922829/c
LOCUS CA922829 812 bp mRNA linear EST 09-MAY-2003
DEFINITION EST040547 MTUS Medicago truncatula cDNA clone MTUS-58H7, mRNA
sequence.
ACCESSION CA922829
VERSION CA922829.1 GI:27409759
KEYWORDS EST.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 812)
VandenBosch, K., Endre, G., Silverstein, K., Town, C.D., Van Aken, S.,
Utterback, T., Cheung, F. and Fraser, C.M.
The Medicago truncatula 6K unigene set: cDNA clones selected and
re-arrayed from various libraries
Unpublished (2002)
Contact: VandenBosch K
Department of Plant Biology
University of Minnesota
220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
Tel: 612 624 2755
Fax: 612 625 1738
Email: kvandenb@cs.umn.edu
Alias clone pMHRP-28F3
TIGR sequence name: MTUCF91TV
More information is available at: www.medicago.org
Seq primer: (gtA AtA CgA Ctc Act AtA ggg C).
Location/Qualifiers
1..812
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="A17"

FEATURES
source
```

/db_xref="taxon:3880"
/clone="MTUS-58H7"
/tissue_type="mixed tissues"
/dev_stage="various stages"
/lab_host="XL0LR"
/clone_lib="MTUS"

/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XL0LR cells."

ORIGIN

Query Match 30.6%; Score 249.6; DB 14; Length 812;
Best Local Similarity 70.6%; Pred. No. 5e-42;
Matches 333; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY 199 TTTTACCTGCTGTTAGTGGCCAAATATCTCTGATATATTCAGACTCGAAATTTTACA 258
DB 623 TTCTTCCCAAAATCAGAGCTCAAAACCTTCTGACTATGTTCTGAATCCAAAGTTTACA 564

QY 259 AGTGAAGCAATTTGACACATGAGATCCAGCAAGTTTTCATCGGCTTTACTCAAAA 318
DB 563 AAGTTGAAGCCATTCTCAGGCCATGAGAAATCCCTCAGGTTTCTTCGGGTTTGTGAAA 504

QY 319 TCGGATTCGAGGTGTACTGTTCTGATGTCAGAGGGTTTGGTCACAAAGAGGTTCTA 378
DB 503 TCGGAATTCGTGGTGCATGATCTGATGTCAGAGGGTTTGGTCTCAGGGTGGCTCAA 444

QY 379 CCGAGAGACACCGTGGCTCTGAGTTCTCGAAGACAAATTTGTTCTAAAGTTAAGATCG 438
DB 443 AAGAGAGGCGAGGAGGCTCGAATTTCTGAGACAAATTTGTTGCCAAGTTTAAATGG 384

QY 439 AATCTGTTTAAAGAACCAAGTGAATCTGATCAACAAATTTGAAGAGGACAA 498
DB 383 AATAGTGGTGAGAAAGACCGAGTTGAGGAGTGAATTAACAAATTTAGAGACGCCAA 324

QY 499 CGACAGAGAGATTTGTGTGGCAGAAATTTTGTGCTGTCAGATGTCATAAGAG 558
DB 323 GAATCGGGAGATTTGTGTGGCAGAAATTTTGTGATCCCTGATCTGATGAATAGAA 264

QY 559 TTAGGACAGTGGCGTGGGAGAAAGCAGAGAGATGACTGGTGATATCCTTCACCGT 618
DB 263 TCCGCAAGTGGCGTGGGAGACGCTGAGAGGATGGCTGGGGGACTAACTGACCGT 204

QY 619 CTTAGGAACAAACAGAGCTCAAGATGGTTTTTTTTTTTTCATTTTCGGTCT 670
DB 203 TGTCTGTTTGAACAAAGATTAGCAATTTGCAATTTGCTCTCTCTCTCTTT 152

RESULT 12

BI311072 714 bp mRNA linear EST 20-JUL-2001
LOCUS EST5312822 GESD Medicago truncatula cDNA clone pGESD9J4 5' end,
DEFINITION mRNA sequence.
ACCESSION BI311072.1 GI:14985399
VERSION BI311072
KEYWORDS EST
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 714)
Grusak, M.A., Samac, D.A., Town, C.D., Van Aken, S., Utterback, T.,
Cho, J. and Fraser, C.M.
ESTs from developing reproductive tissues of Medicago truncatula
Unpublished (2001)
Contact: Michael A. Grusak
USDA/ARS Children's Nutrition Research Center

REFERENCE

AUTHORS
TITLE
JOURNAL
COMMENT

Baylor College of Medicine
1100 Bates Street, Houston, TX 77030-2600, USA
Tel: 713-798-7044
Fax: 713-798-7078
Email: mgrusak@bcm.tmc.edu
B397927e

TIGR sequence name: MTPAP50TK
More information is available at: www.medicago.org
Seq primer: SKmod (CTA GAA CTA gtg GAT CC).

FEATURES

source

1. 714
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="A17"
/db_xref="taxon:3880"
/clone="pGESD9J4"
/tissue_type="immature seeds"
/dev_stage="immature seeds, 11 to 19 days after
pollination"
/clone_lib="GESD"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; Immature seeds, collected from pods ranging in age
from 11 to 19 days after pollination, were harvested from
greenhouse-grown plants. Seeds were removed and
separated from pod walls and were immediately frozen in
liquid nitrogen. Seeds throughout the age range were
pooled for mRNA extraction. cDNA was prepared from polyA+
enriched RNA. The cDNA was directionally ligated into
the Unizap XR vector from Stratagene and packaged using
Gigapack III Gold packaging extracts. Plasmids containing
cDNA inserts were excised from the recombinant lambda-Zap
phage using Ex-assist helper phage and propagated in
XL0LR cells."

ORIGIN

Query Match 30.3%; Score 247.8; DB 12; Length 714;
Best Local Similarity 67.6%; Pred. No. 1.2e-41;
Matches 348; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 199 TTTTACCTGCTGTTAGTGGCCAAATATCTCTGATATATTCAGACTCGAAATTTTACA 258
DB 172 TTCTTCCCAAAATCAGAGCTCAAAACCTTCTGACTATGTTCTGAATCCAAAGTTTACA 231

QY 259 AGTGAAGCAATTTGACACATGAGATCCAGCAAGTTTTCATCGGCTTTACTCAAAA 318
DB 232 AAGTTGAAGCCATTCTCAGGCCATGAGAAATCCCTCAGGTTTCTTCGGGTTTGTGAAA 291

QY 319 TCGGATTCGAGGTGTACTGTTCTGATGTCAGAGGGTTTGGTCACAAAGAGGTTCTA 378
DB 292 TGGGAATTCGTGGTGCATGATCTGATGTCAGAGGGTTTGGTCTCAGGGTGGCTCAA 351

QY 379 CCGAGAGACACCGTGGCTCTGAGTTCTCGGAAGACAAATTTGTTGCTAAAGTTAAGATGG 438
DB 352 AAGAGAGCGAGGAGGCTCGAATTTCTGAGACAAATTTGTTGCCAAGTTTAAATGG 411

QY 439 AAATCGTTTGAAGAAAGACCAAGTGAATCTGATTAATCAACACATTAATTAAGAGCAA 498
DB 412 AAATAGTGGTGAGAAAGACCGAGTTGAGGAGTGAATAAACAATAATTAAGAGACGGCA 471

QY 499 GGACAGAGAGATTTGGTGTGCGCAAGATTTTGTGCTGCTGATGTCATGATGATGAG 558
DB 472 GAATCGGGAGATTTGGTGTGCGCAAAATTTTCTGATCCCTGATCTGATGATTAAGAA 531

QY 559 TTAGGACAGTGGCGTGGGAGAAAGCAGAGAGATGACTGGTGATATGCTTTCACCGT 618
DB 532 TCCGCAAGTGGCGTGGGAGACGCTGAGAGGATGGCTGGGGGACTAACTGACCGT 591

QY 619 CTTAGGAACAAACAGAGCTCAAGATGGTTTTTTTTTTTTCATTTTCGGTCTCTAGATTC 678
DB 592 TGTCTGTTTGAACAAAGATTAGCAATTTGCAATTTGCTCTCTCTCTCTCTTTCCCTGTA 651

QY 679 TCGGAATTAATGATGAGTCTGCTGTTGCTTT 713

```

Db      652 GAATTGATACAAAGATGATGTTGTGTGTTATTTT 686

RESULT 13
LOCUS   BE239550
DEFINITION EST403599 MHRP- Medicago truncatula cDNA clone pMHRP-28F3, mRNA
sequence.
ACCESSION BE239550
VERSION   BE239550
KEYWORDS  EST
SOURCE    Medicago truncatula (barrel medic)
ORGANISM  Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE
1 (bases 1 to 597)
Harrison, M.J., Town, C.D., Bowman, C.L., Craven, M.B., Hansen, T.S.,
Holt, I.E., Cho, J. and Fraser, C.M.
ESTs from phosphate-starved roots of Medicago truncatula
Unpublished (2000).
Contact: Maria J. Harrison
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73401, USA
Tel: 580-223-5810
Fax: 580-221-7380
Email: mjharrison@noble.org
The Samuel Noble Roberts Foundation: N265417e
TIGR sequence name: MTHAC26TK
More information is available at:
http://chrysis.tamu.edu/medicago
Seq primer: SKmod (CTA gAA CTA gTG gAT CC).
FEATURES             source
1..597
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="A17"
/db_xref="taxon:3880"
/clone="pMHRP-28F3"
/tissue_type="roots"
/dev_stage="phosphate-starved"
/lab_host="XLOLR"
/clone_lib="MHRP-"
/notes="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; At the trifoliolate stage, M. truncatula plants were
transplanted to phosphate-free sand and grown for a
further 30 days. During this period, they were fertilized
twice weekly with 1/2 Hoaglands solutions containing 200mM
potassium phosphate. cDNA was prepared from polyA+
enriched RNA. The cDNA was directionally ligated into the
Unizap XR vector from Stratagene and packaged using
Gigapack III Gold packaging extracts. Plasmids containing
cDNA inserts were excised from the recombinant lambda-Zap
phage using Ex-assist helper phage and propagated in
XLOLR cells."

ORIGIN
Query Match      30.2%; Score 246.6; DB 10; Length 597;
Best Local Similarity 75.6%; Pred. No. 2.4e-41;
Matches 306; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY      199 TTTTACCTGCTGTAGTCCCAATATCTCTGATTATATCCAGACTCGAAATTTTACA 258
Db      173 TTCTTCCAAATCAGAGCTCAAAACCTTCTGATGATGTTCTGATCCAGTTTACA 232
QY      259 AGGTGGAAGCAATGTCCAGACCATCGAGAAATCCAGCAAGTTTATCGCTTTACTGAAA 318
Db      233 AAGTTGAAGCAATTTCTCAGGCCATCGAGAAATCCCTCAGGTTTCTTCGGGTTTGTGAAA 292
QY      319 TCGGGATTCAGGTTACTGTTTCTGATGTGAGAGGTTTGGTGCACAGGAGTTCTA 378

```

```

Db      293 TGGGAATTGTTGGTCTCACTGATCTGATGTCAAGGGGTTTGGTCTCAGGGTGGCTCAA 352
QY      379 CCGAGAGACACGGTGGCTCTGAGTTCTCGAAGACAAAATTTGTTGTTAAAGTTAAGTTGG 438
Db      353 AAGAGAGGCGAGGAGGCTCCGAATTTCTGAACACAAATTTGTTGCGAAAGTTAAATGG 412
QY      439 AAATCGTTCTTAAAGAGACCAAGTGGATCTCTATCAACACAAATTAATGAAGGACAA 498
Db      413 AAATAGTGTGTGAGAAAGACCAAGGTTGAGCAGTGATTAACAAATTTATCGAGACGSCAA 472
QY      499 GGACAGGAGAGATTGGTGATGCAAGATTTTGTGTCCTGTGTGATGTGTATGATGAG 558
Db      473 GAATCGGGAGATTGGTGATGCGAAATTTCTTGATCCCTGATCTGATGATTAAGAA 532
QY      559 TTAGGACAGGTGAGCTGGGGAGAAAAGACGAGAAAGATGACTGGTG 603
Db      533 TCCGCACAGTGTGAGCGTGGGAGCAGGCTGAGAGGATGGCTGGGG 577

RESULT 14
LOCUS   AW035791
DEFINITION EST281945 tomato callus, TAMU Lycopersicon esculentum cDNA clone
sequence.
ACCESSION AW035791
VERSION   AW035791
KEYWORDS  EST
SOURCE    Lycopersicon esculentum (tomato)
ORGANISM  Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 613)
Alcala, J., Vrebalov, J., White, R., Matern, A.L., Vision, T.,
Holt, I.E., Liang, F., Upton, J., Craven, M.B., Bowman, C.L., Ahn, S.,
Ronning, C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and
Giovannoni, J.
Generation of ESTs from tomato callus tissue
Unpublished (1999)
Contact: CUGI
Clemson University
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
FEATURES             source
1..613
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA96"
/db_xref="taxon:4081"
/clone="cLEC36D8"
/tissue_type="callus"
/dev_stage="25-40 days old"
/lab_host="XL1-Blue MRF"
/clone_lib="tomato callus, TAMU"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Giovannoni laboratory; cLEC - Cotyledons
of seedlings 7-10 days post-germination were excised, cut
at both ends and placed on MS medium with no selection.
Mixed callus was harvested at 25 and 40 days and included
undifferentiated masses. Tomato Callus EST Library"

ORIGIN
Query Match      30.0%; Score 245; DB 9; Length 613;
Best Local Similarity 72.5%; Pred. No. 5.1e-41;
Matches 317; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY      167 GGTCAAAAGTACCGAGTATAACAGTGTGTTTACCTGTCTGTTAGTCCCAATATC 226
Db      131 GCTCACCTTCAAAACGCTGTCAAAATGTCTCTTCTCCGATTATCAGAGCCCAAACTC 190
QY      227 TTCTGATTATTCCAGACTCGAAATTTTACAAGGTGGAAGCAATTTGTGACACCATGGAG 286

```

```

191 TCAGATTTTCGTCCTGATGCGAAGTTTACAAAGTTGAAGCAATCTAAGACCTTGGAG 250
287 AATCCAGCAAGTTTCAATCGGCTTACTGAAATCGGATTCGAGGTGTTACTGTTCTGA 346
251 AATCAACAGTTTCTTGGGACATCTCTGAAATGCGCATTCGTGTTCTACTGTTTGGGA 310
347 TGTGAGAGGTTTGTGTCACAAAGAGTTCTACCGAGAGACACGGTGGCTCTGAGTTCTC 406
311 TGTGTTGTTTGTGGCGCCCAAGTGGCTTGAATGAGAGGCAAGCTGGCTCTGAATCTC 370
407 GGAAGCAAAATTTGTTGCTTAAGTTAAGTGAATCGTTGTTAAGAAACCAAGTGA 466
371 TGAAGCACGTTTGTGTCAAAGTTAAATGGAATTTGTTGTCAGCAAGACAGAGTTGA 430
467 ATCTGTAATCAACACAATTAATGAAGGCAAGGACAGAGAGATTGTTGATGCGAAGAT 526
431 AGAGTCATTTGCCATGATTAATGAGAGGCAAGACTGGTGAATAGTTGTAAGAT 490
527 TTTTGTGTTTGCCTGTTGTCAGATGTCATPAGAGTTAGACAGGTGAGCGTGGGAGAAAGC 586
491 ATTCTTGACTCCCATCTCCGATGTTATPAGAGTTGCGACTGGTGAACCGGAGAAAGGC 550
587 AGAAGAGTACCTGGTG 603
551 AGAGAGGATGATGGGAG 567

```

```

RESULT 15
BI932123
LOCUS
DEFINITION
EST552012 tomato flower, 8 mm to preanthesis buds Lycopersicon
esculentum cDNA clone cTOC22G23 5' end, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Lycopersicon esculentum (tomato)
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 702)
van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
Utterback,T., Van Aken,S., Rönning,C.M., Nierman,W., Fraser,C.M.,
Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
Generation of ESTs from tomato flower tissue, buds 8 mm -
preanthesis
Unpublished (2001)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
This clone is available through the Clemson University Genomics
Institute
Seq primer: T3.
Location/Qualifiers
1..702
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cTOC22G23"
/tissue_type="flower"
/dev_stage="buds 8mm to preanthesis"
/clone_lib="tomato flower, 8 mm to preanthesis buds"
/note="Vector: pBluescript SK(-); Site1: EcoRI; Site2:
XhoI; supplier: Cornell University; sequencing: The
Institute for Genomic Research; Flower buds and flowers
were taken from greenhouse plants (4-8 wks old, TA496).
They were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

```

```

FEATURES
source
1..702
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cTOC22G23"
/tissue_type="flower"
/dev_stage="buds 8mm to preanthesis"
/clone_lib="tomato flower, 8 mm to preanthesis buds"
/note="Vector: pBluescript SK(-); Site1: EcoRI; Site2:
XhoI; supplier: Cornell University; sequencing: The
Institute for Genomic Research; Flower buds and flowers
were taken from greenhouse plants (4-8 wks old, TA496).
They were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

```

ORIGIN

```

Query Match 30.8%; Score 245; DB 12; Length 702;
Best local similarity 72.5%; Pred. No. 4.8e-41;
Matches 317; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 167 GGTCAACAAGTCCCGAGTAATACAGTCGTGTTTACCTGTGCTTGTAGTGCCCAATATC 226
Db 172 GTCACCTTCAACAGCTGTCAAAATGCTCTTCTTCCCGATTATCAGAGCCCAAACTC 231
QY 227 TTCTGATTATATTCAGAGACTCGAAATTTTACAAGGTGGAAGCAATGTCTAGACCATGGAG 286
Db 232 TCCAGATTTTGTGCTGATCGAAGTTTACAAAGTTGAAGCAATCTAAGACCTTGGAG 291
QY 287 AATCCAGCAAGTTTCAATCGGCTTACTGAAATCGGATTCGAGGTGTTACTGTTCTGA 346
Db 292 AATTCACAGGTTTCTTCCGCACTACTGAAATGCGCAATTCGTGGTGTCTGTTTCGA 351
QY 347 TGTGAGAGGTTTGTGTCACAAGGAGGTTCTACCGAGAGACACGGTGGCTCTGAGTTCTC 406
Db 352 TGTGTTGTTTGTGGCGCCCAAGTGGCTTGAATGAGAGGCAAGCTGGCTCTGAATCTC 411
QY 407 GGAAGCAAAATTTGTTGCTTAAAGTTAAGATGGAATCGTTGTTAAGAAACCAAGTGA 466
Db 412 TGAAGCACGTTTGTGTCAAAGTTAAATGGAATTTGTTGTCAGCAAGACCAAGTTGA 471
QY 467 ATCTGTAATCAACACAATTAATGAGAGGCAAGGACAGAGAGATTGTTGATGCGAAGAT 526
Db 472 AGGAGTCATTTGCCATGATTAATGAGAGGCAAGCAATTCGTGTAATAGTGAAGAT 531
QY 527 TTTTGTGTTTGCCTGTTGTCAGATGTCATPAGAGTTAGACAGGTGAGCGTGGGAGAAAGC 586
Db 532 ATTCTTGACTCCCATCTCCGATGTTATPAGAGTTGCGACTGGTGAACCGGAGAAAGGC 591
QY 587 AGAAGAGTACCTGGTG 603
Db 592 AGAGAGGATGATGGGAG 608

```

Search completed: May 28, 2004, 00:14:51
Job time : 1898.7 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2004, 13:35:34 ; Search time 2651.35 Seconds
(without alignments)
14663.718 Million cell updates/sec

Title: US-09-756-541-14
Perfect score: 897
Sequence: 1 GGGGTGCGCCGCTCTAGA.....AAAAAATAAACTCGAG 897

Scoring table: IDENTITY NUC
Gapop 1c_0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg_hum.*

31: em.htg_inv.*

32: em.htg_other.*

33: em.htg_mus.*

34: em.htg_pin.*

35: em.htg_rsd.*

36: em.htg_mar.*

37: em.htg_vit.*

38: em_sy.*

39: em.htgo_hum.*

40: em.htgo_mus.*

41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	897	100.0	897	6	AR125590	Sequence
2	840	93.6	840	8	AF095454	AR125590 Sequence
3	594	66.2	594	6	AR125592	AR125592 Sequence
4	300.2	33.5	796	8	AY027892	Medicago
5	286.4	31.9	817	6	AR125589	Sequence
6	286.4	31.9	844	8	AF095455	Arabidops
7	285	31.8	588	6	AR125591	Sequence
8	285	31.8	591	6	AX507438	Sequence
9	277.4	30.9	947	8	BT005209	Arabidops
10	277.4	30.9	947	8	AY442185	Lycopersi
11	217.2	24.2	902	8	AK068407	Oryza sat
12	217.2	24.2	902	8	AK099152	Oryza sat
13	215.4	24.0	1085	8	PF1489604	pinus pin
14	107	11.9	686	1	AF017419	Nostoc pu
15	103.2	11.5	2493	1	ASP251822	Anabaena
16	103.2	11.5	341880	1	AP003589	Nostoc sp
17	99	11.0	384	1	FDGLNBPRT	X97327 F.diplosiph
18	98.8	11.0	191028	8	DPUS8804	U38804 Porphyra pu
19	96.4	10.7	11978	1	U67574	Methanococc
20	96.4	10.7	110000	6	AR271569_01	Continuation 12 of
21	96.2	10.7	10338	1	U67464	U67464 Methanococc
22	96.2	10.7	110000	6	AR271569_14	Continuation 15 o
23	95.4	10.6	383	1	SYOGLNB_137	M62447 Synethococc
24	95.4	10.6	4348	1	AP079137	AF079137 Synethococ
25	94.4	10.5	452	1	SSGLNBP	X97496 Synethococyst
26	94.4	10.5	130001	1	D90915	D90915 Synethococyst
27	92	10.3	300143	1	AE017165	AE017165 Prochloro
28	90.4	10.1	164921	8	AF022186	AP005370 Thermosyn
29	90.4	10.1	299350	1	AP005370	AP005370 Thermosyn
30	88.4	9.9	9973	1	AE004295	AE004295 Vibrio ch
31	88.2	9.8	1622	1	AF120107	AF120107 Syrachoco
32	86.8	9.7	300045	1	AE016803	AE016803 Vibrio vu
33	86.8	9.7	302320	1	AP005081	AP005081 Vibrio pa
34	86.4	9.6	262202	1	EX572094	EX572094 Prochloro
35	85.2	9.5	247950	1	AP005340	AP005340 Vibrio vu
36	84.8	9.5	96109	6	AK08756	AK08756 Sequence
37	84.8	9.5	96109	6	AX067460	AX067460 Sequence
38	84	9.4	11371	1	AE006236	AE006236 Pasteurel
39	83.6	9.3	375	6	AR318761	AR318761 Sequence
40	83.6	9.3	298950	1	AP004597	AP004597 Oceanobac
41	83.4	9.3	9977	1	AE002548	AE002548 Neisseria
42	83.2	9.3	300880	1	AE016917	AE016917 Chromobac
43	82.2	9.2	300272	1	AE017213	AE017213 Geobacter
44	82	9.1	346792	1	EX571658	EX571658 Wolinella
45	81.8	9.1	349061	1	NMA222491	ALI62753 Neisseria

ALIGNMENTS

RESULT 1
AR125590
LOCUS AR125590 Sequence 14 from patent US 6177275.
DEFINITION AR125590 Sequence 14 from patent US 6177275.
ACCESSION AR125590
VERSION AR125590.1 GI:14111652
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 897)
AUTHORS Coruzzi,G.M., Lam,H.-M. and Hsieh,M.-H.
TITLE Plant nitrogen regulatory P-PII genes
JOURNAL Patent: US 6177275-A 14 23-JAN-2001;
FEATURES Location/Qualifiers

Pred. No. is the number of results predicted by chance to have a

source	1..897	100.0%	Score 897;	DB 6;	Length 897;		
ORIGIN	Query Match	Best Local Similarity	Matches 897;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	CGCGTGTGCGCGCTCTAGAACTAGTGGATCCCCGGGCTGCAGGAATTCGGCACGAGGC	60				
Db	1	CGCGTGTGCGCGCTCTAGAACTAGTGGATCCCCGGGCTGCAGGAATTCGGCACGAGGC	60				
Qy	61	TACTCGGAAACTGGGCTTGCTCACTCCCTTCATCTTAATAACATCAAGAAAGAAATCCC	120				
Db	61	TACTCGGAAACTGGGCTTGCTCACTCCCTTCATCTTAATAACATCAAGAAAGAAATCCC	120				
Qy	121	TGTTTTGATTTTCAGTTTGTGTTTTGTCACAGACTTAGACATTTCTCGGTTTTCTCATTTTAA	180				
Db	121	TGTTTTGATTTTCAGTTTGTGTTTTGTCACAGACTTAGACATTTCTCGGTTTTCTCATTTTAA	180				
Qy	181	CACGCGGTCAGAGCGGTAAAGATATGCCCGCTCGTTCCTGTGATTAATGCCCAAGATC	240				
Db	181	CACGCGGTCAGAGCGGTAAAGATATGCCCGCTCGTTCCTGTGATTAATGCCCAAGATC	240				
Qy	241	GCCTGACTCAATTCCTCATGCTAAATTTCTACAAAGTGGAAAGCAATTTCTCAGGCCCTGCGC	300				
Db	241	GCCTGACTCAATTCCTCATGCTAAATTTCTACAAAGTGGAAAGCAATTTCTCAGGCCCTGCGC	300				
Qy	301	AGTCTCGAAGTTTCCTCGGCTTTGCTTAAAAATTTGGTATTGAGGTGTTCATGTTTCTGA	360				
Db	301	AGTCTCGAAGTTTCCTCGGCTTTGCTTAAAAATTTGGTATTGAGGTGTTCATGTTTCTGA	360				
Qy	361	TGTTTCGAGGTTTTCGCTCAAGGTGGTTCAACTGAGAGGCAGGCGGCTCAGAAATTTTC	420				
Db	361	TGTTTCGAGGTTTTCGCTCAAGGTGGTTCAACTGAGAGGCAGGCGGCTCAGAAATTTTC	420				
Qy	421	TGAAGACAAGTTTGTCTCTAAAGTTAAAGATCGAGATCGTGTAGCAAGAACCAGTTTGA	480				
Db	421	TGAAGACAAGTTTGTCTCTAAAGTTAAAGATCGAGATCGTGTAGCAAGAACCAGTTTGA	480				
Qy	481	GGATGTTATAGAAAATTCATTGAGGAGGCAAGAACTGGAGAGATTCGAGAGCGCAAGAT	540				
Db	481	GGATGTTATAGAAAATTCATTGAGGAGGCAAGAACTGGAGAGATTCGAGAGCGCAAGAT	540				
Qy	541	TTTCTTCTGCTCTTTTCAGATGTAATAAGAGTCCGCACTGGTCAGCGGGGTGATAAGGC	600				
Db	541	TTTCTTCTGCTCTTTTCAGATGTAATAAGAGTCCGCACTGGTCAGCGGGGTGATAAGGC	600				
Qy	601	TGAGAGATGACAGAGGCGGATCTGACATGATGACTTCTGCTTGA CTGCTGTGACCCAGC	660				
Db	601	TGAGAGATGACAGAGGCGGATCTGACATGATGACTTCTGCTTGA CTGCTGTGACCCAGC	660				
Qy	661	AATATAGCAATTCAGACTTAACTGTCTTTGAGAAAGCCCGCCCTTATTAGCCATTATCC	720				
Db	661	AATATAGCAATTCAGACTTAACTGTCTTTGAGAAAGCCCGCCCTTATTAGCCATTATCC	720				
Qy	721	AGTATAGCTTGATTAATTTGAAATTTTTGTTTTCTTTAACTAAAGAAAACAAAGATCTTTTCA	780				
Db	721	AGTATAGCTTGATTAATTTGAAATTTTTGTTTTCTTTAACTAAAGAAAACAAAGATCTTTTCA	780				
Qy	781	TTATCTCTTGATGATTAATTCGAAACGGAAGGATCGCGAAATTTGTTCAAGTCTTTCGAAG	840				
Db	781	TTATCTCTTGATGATTAATTCGAAACGGAAGGATCGCGAAATTTGTTCAAGTCTTTCGAAG	840				
Qy	841	ATAAATAACAGAGAGAGGATTAATGTTTAAACAAAAAATAAAAAAATAAAAACTCGAG	897				
Db	841	ATAAATAACAGAGAGGAGTAATGTTTAAACAAAAAATAAAAAAATAAAAACTCGAG	897				
RESULT 2							
AF095454							
LOCUS							

RESULT 2
AF095454
LOCUS

Db 127 CTCGATTGGTCTCAAAAGTCACCGAGTAATAACAGTCGTGTTTACCTGCTGTAGTSCC 186
 QY 233 CAAAGCTCGCTGACTACATCTCTGATGCTAAATTTCTACAAAGTGGAGCAATTTCTCAGG 292
 Db 187 CAATATCTCTGATATATTCAGACTCGAAATTTTACAGGTGGAGCAATTTGTCAGA 246
 QY 293 CCTGCGAGTCTCGCAAGTTTCTCGGCTTCTGCTAAATTTGCTATGATGAGGTTTACT 352
 Db 247 CCATGAGATCCAGCAAGTTTCATCGCTTTTACTGAAATTCGGATTCGAGGTGTTACT 306
 QY 353 GTTCTGATGTTGAGGTTTGGTCTCAAGTGTCTCACTGAGAGGCGGCGGCTCA 412
 Db 307 GTTCTGATGTTGAGGTTTGGTCAAGAGGTTCTACCGAGAGACAGCGTGGCTCT 366
 QY 413 GAATTTTCTGAAGCAAGTTTGTGCTAAAGTTAAGATGGAGTCGTGTTAGCAAGAC 472
 Db 367 GAGTCTCGGAAGCAATTTGTGCTAAGTTAAGATGGAAATCGTTTAAAGAGAC 426
 QY 473 CAGGTTGAGATGTTATAGAAAAATCATTTGAGAGGCAAGAACTGGAGATTTGGAGAC 532
 Db 427 CAAGTGGAACTCTGTAATCAACAAATAATGAAGGACCAAGGACAGGAGATTTGGTAT 486
 QY 533 GGCAGATTTTCTGCTGCTGTTTCAAGTCTAATAGATCCGACTGTTGAGCGGGT 592
 Db 487 GGCAGATTTTGTGTTTTCCTGCTGCTCATAGATGTCATAGAGTTAGACAGGTGAGCGTGG 546
 QY 593 GATAAGGCTGAGAGGATGACAGGAG 617
 Db 547 GAGAAAGCAGAGAGATGACTGGTG 571

RESULT 9

B7005209 591 bp mRNA linear PLN 14-MAR-2003
 LOCUS Arabidopsis thaliana At4g01900 mRNA, complete cds.
 DEFINITION B7005209
 ACCESSION B7005209.1 GI:28950698
 VERSION FLI CDNA.
 KEYWORDS Arabidopsis thaliana (thale cress)
 SOURCE Arabidopsis thaliana

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 591)
 Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Carninci, P.,
 Chan, M.M., Chang, C.H., Dale, J.M., Hayashizaki, Y., Hsuan, V.W.,
 Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
 Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,
 Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,
 Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K.,
 Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and
 Ecker, J.R.

REFERENCE

Arabidopsis ORF clones
 Unpublished
 2 (bases 1 to 591)
 Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Carninci, P.,
 Chan, M.M., Chang, C.H., Dale, J.M., Hayashizaki, Y., Hsuan, V.W.,
 Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
 Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,
 Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,
 Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K.,
 Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and
 Ecker, J.R.

TITLE

Journal Submission
 Submitted (14-MAR-2003) Salk Institute Genomic Analysis Laboratory
 (SIGAL), Plant Biology Laboratory, The Salk Institute for
 Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
 USA

JOURNAL

RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN
 Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J.,
 Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
 Hayashizaki, Y. and Shinozaki, K.

COMMENT

The Salk, Stanford, PGSC (SSP) Consortium members constructed and
 sequenced the pUNI (ORF) clones using the RAPL cDNAs: Kim, C.J.,
 Chen, H., Cheuk, R., Shinn, P., Bowser, L., Chan, M.M., Chang, C.M.,
 Dale, J.M., Hsuan, V.W., Jones, T., Karlin-Neumann, G., Lam, B.,
 Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J.,
 Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C.,
 Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A., and
 Ecker, J.R.

Kim, C.J. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
 this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
 contributed equally to this work as PIs.

FEATURES

source

1. 591
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="4"
 /clone="U23463"
 /ecotype="Columbia"
 /note="This clone is in pUNI 51"

CDS

1. 591
 /note="p II nitrogen sensing protein GLB I"
 /codon_start=1
 /product="At4g01900"
 /protein_id="AA063273.1"
 /db_xref="GI:28950699"
 /translation="MAASMTKPISTISLGFYSDRNIAFSDICISCFHRSRSCLD
 LVTKPSNNRVLVPSAQISSDIYDPSKFYKVAIVPMRIQOVSSALLKIGINGVT
 VSDVGFCAQGSSTERHSGSEFSEDFKFAKMEIVKKQDQVESVINTIIGARTGEI
 GDGKIFVLVPSDVRVTRTGERGERKAEKWTGMDLSPS"

ORIGIN

Query Match 31.8%; Score 285; DB 8; Length 591;
 Best Local Similarity 70.3%; Pred. No. le-59;
 Matches 397; Conservative 0; Mismatches 165; Indels 3; Gaps 1;
 QY 56 GAGGCTACTGGGAACTGGGCTTGCCTCACTCTCTTCAATCTTAATAACATCAAGAAGAA 115
 Db 7 GCGTCAATGACGAAACCCATCTCAATAACTTCTCTCGGTTCTTATCTGATCGAAGAAC 66
 QY 116 TTCCTGTTTTTGAATTCAGTTTGTGTCAGAGCTTAGACATTCCTCGGTTTCTCTAC 175
 Db 67 ATGCTTCTCTGATTCGATTCGATTTGTTCTGATTCAGATTCGACCATCTTCG 126
 QY 176 TTAAACACCGCGCTCAAGCGGTAAGATATGCC---CCGTCGTTCTGTGATTAATGCC 232
 Db 127 CTCGATTTGGTCACAAAGTCACCGAGTAATAACAGTCGTGTTTACCTGCTGTAGTGC 186
 QY 233 CAAAGCTCGCTGACTACATTCCTGATGCTAAATTTCTACAAAGTGAAGCAATTTCTCAGS 292
 Db 187 CAAATATCTCTGATTTATATTCGACTCGAAATTTTACAGGTGGAACAATTTGCAGA 246
 QY 293 CCCTGCGAGTCTCGCAAGTTTCTCGGCTTTGCTTAAATTTGGTATTCGAGGTGTACT 352
 Db 247 CCATGAGAAATCCAGCAAGTTTCAICGGCTTTACTCGAAATCGGATTCGAGGTGTACT 306
 QY 353 GTTCTGATGTTTCTGAGGTTTGTGCTCAAGGTGTTTCACTGAGAGGCGAGCGGCTCA 412
 Db 307 GTTCTGATGTTGAGAGGGTTTGTGTCACAAAGAGGTTTCTACCGAGAGACACCGTGGCTCT 366
 QY 413 GAATTTTCTGAAGACAAAGTTTGTGCTAAAGTTAGATGAGATCGTGGTTAGCAAGAC 472
 Db 367 GAGTTCTCGAAGACAAATTTGTTGCTAAAGTTAAGATCGAATCGTTTGAAGAAGAC 426
 QY 473 CAGGTTGAGATGTTATAGAAAAATCATTTGAGAGGCAAGAACTCGAGAGATTTGAGAC 532
 Db 427 CAAAGTGGAAATCTGTAATCAACAATTAATTGAAGGCAAGGACAGAGATTTGAT 486
 QY 533 GGCAGATTTTCTGCTGCTGTTTTCAGATGTAATAAGAGTCCGCACTGTTGAGCGGGGT 592
 Db 487 GGCAGATTTTGTGTTTTCCTGCTGTCATAGATGTTAGGACAGGTGAGCGTGGG 546

Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@ias.affrc.go.jp)
Tel:81-29-838-7007, Fax:81-29-838-7007

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, R., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariikawa, R., Nishikawa, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaka, R., Ohno, M., Otsu, N., Ota, Y., Saichoh, H., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, S., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers

1. 902
/organism="Oryza sativa (japonica cultivar-group)"
/mol type="mRNA"
/cultivar="Nipponbare"
/db xref="taxon:39947"
/clone="J013149B08"

ORIGIN

Query Match 24.2%; Score 217.2; DB 8; Length 902;
Best Local Similarity 69.7%; Pred. No. 6.3e-43;
Matches 294; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

Qy 238 CTGCGCTGACTACATTCCTGATGCTAAATCTCAAAAGTGAAGCAATTCACGCCCTG 297
Db 267 CGCAGCAGGGTACCAGCGGAGTCGGAGTCTTCAAGGTGGAGGCAATCTCTGAGGCCATG 326
Qy 298 GCGAGTCTCGCAAGTTCTCTCGCTTTGCTAAATTTGTTATTCGAGGTGTACTGTTC 357
Db 327 GAGGTCCTTATGTGTCTATCGGTTTCTGCAATGGGATCAGGCGTGCAGCGTGC 386
Qy 358 TGATGTTTCGAGGTTTTCGTTGCTCAAGGTGTTTCAACTGAGAGGCGGCGCTCAGAATT 417
Db 387 CGAGCTGCGGGTTTTCGCGGCACAGGCGGGTCAACTGAGAGGCGATGAAGGTCAGAATT 446
Qy 418 TCTGAGACAGAGTTTCTGATTAAGTAAAGTGAAGTGGATCGGTAGCAAGACCGAT 477
Db 447 TGCAGAGATACATTTATGATAAAGTTAAGATGAAATAGTGGTTCCTCAAGGATCAGGT 506
Qy 478 TGAGGATGTTATAGAAAATTCATTGAGGAGCGAAGCACTGAGAGATTGGAGCGGCA 537
Db 507 TGAAGCTGTTGTCACAGATAATTGAAAGGCAAGAACAGAGAAATTTGATGTGAA 566
Qy 538 GATTTCTGCTGCTGTTTCGATGTAATGAAGTCCGCACTCGTGAAGCGGGTGATA 597
Db 567 AATATTTTGTATCCCGTTCGGACGTGATCAGAATAAGCAAGCGGCAAGCGGAGCG 626
Qy 598 GCGTGAAGATGACAGAGGCGGATCTGATGATGATGATCTCTCTGATGCTGTGAC 657
Db 627 AGCGAGAGGATGCGCGGAGGCTGGCGGACAAAGCTGTCTCTCAGCAATGCGGATCTCATG 686
Qy 658 AG 659

Db 687 AG 688

RESULT 12

AK099152
LOCUS
DEFINITION
ORIGIN
ACCESSION
KEYWORDS
SOURCE

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

ORIGIN

Isikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

PAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Tshibiki, J., Kawanata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, K., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Mamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Maeuguma, T., Miyazaki, A., Murata, M., Nakamura, M., Nishik, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

```
FEATURES
source
1. .902
Location/Qualifiers
/organism="Oryza sativa (japonica cultivar-group)"
/moi_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023070H02"
```

ORIGIN

Query Match	24.2%	Score 217.2;	DB 8;	Length 902;
Best Local Similarity	69.7%	Pred. No. 6.3e-43;		
Matches 294;	Conservative	0;	Mismatches 128;	Indels 0;
Gaps 0;				

238	CTCGCCGTGACTACATTCCTGATGCTAAATTTCTACAAAGTCGAGCAATTCCTCAGGCCCTG	297
267	CGCAGCAGGGTACCGACCGGAGTCGGAGTTCTTACAAAGTGGAGCAATCTCTGAGGCCCATG	326
298	GCAGAGTCCTCCCAAGTTTCTCTCGGCTTTGCTTAAAAAATTGGTATTTCGAGGTGTTACTGTGTTTC	357
327	GAGGTCCTTATGTGTTCATCGGGTTTGTCTGCAAAATGGGGNTCAGAGCGCTGACGGTGTC	386
358	TGAGTTCGAGGTTTTGGTGCTCAAGGTGGTTTCAACTGAGAGGCGAGCGCGCTCGAGAATT	417
387	CGACGTGCGGGGTTTTCGCGCACAGGGCGGGTCAACTGAGAGGCGATGAAGGGTTCAGAATT	446
418	TTCTGAAGACAAGTTTCTTGCTTAAAGTTAAAGATGGAGATCGTGGTTAGCAAGACCCAGGT	477
447	TGCAGAGATACATTTATTTCATTAAGTTAAGTATGGAATAGTGGTCTCAAGATCAGGT	506
478	TGAGGATGTTTATAGAAAATTCATTGAGGAGGCGCAAGAACTGGAGAGAGATCGAGACGGCAA	537
507	TGAAGCTGTTGTTTGACAAAGATAATTCGAAAAGGCGAAGAACAGAGCAAAATTCGTGATCGAAA	566
538	GATTTTCTGCTGCCGTGTTTCAGATGTAAATAGAGTCCGCACCTGGTCAGCGGGGTGATAA	597
567	AAATATTTTGTATACCGGTGTCGACGCGTGAATCAGAATACGCACCGCGCGAAACGAGGGGAGCG	626
598	GGCTGAGAGGATGACACGAGGGCGGATCTGCACATGAGTACTTTCGTCTGCACTGTGCTGTGACCC	657
627	AGCGGAGAGATGGCCGAGGGGCTGGCGGACAAGCTGCTCTCAGCAATGCCGATCTCATG	686
658	AG 659	
687	AG 688	

RESULT 13					
PP1489604					
LOCUS	PP1489604	1085 bp	mRNA	linear	PLN 11-OCT-2003
DEFINITION	Pinus pinaster mRNA for pII-like protein (glb gene).				

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE AUTHORS TITLE JOURNAL	REFERENCE AUTHORS TITLE JOURNAL
1. H. J.	2.
3.	4.
5.	6.
7.	8.
9.	10.
11.	12.
13.	14.
15.	16.
17.	18.
19.	20.
21.	22.
23.	24.
25.	26.
27.	28.
29.	30.
31.	32.
33.	34.
35.	36.
37.	38.
39.	40.
41.	42.
43.	44.
45.	46.
47.	48.
49.	50.
51.	52.
53.	54.
55.	56.
57.	58.
59.	60.
61.	62.
63.	64.
65.	66.
67.	68.
69.	70.
71.	72.
73.	74.
75.	76.
77.	78.
79.	80.
81.	82.
83.	84.
85.	86.
87.	88.
89.	90.
91.	92.
93.	94.
95.	96.
97.	98.
99.	100.

FEATURES

gene	CDS
g00001	g00001
g00002	g00002
g00003	g00003
g00004	g00004
g00005	g00005
g00006	g00006
g00007	g00007
g00008	g00008
g00009	g00009
g00010	g00010
g00011	g00011
g00012	g00012
g00013	g00013
g00014	g00014
g00015	g00015
g00016	g00016
g00017	g00017
g00018	g00018
g00019	g00019
g00020	g00020
g00021	g00021
g00022	g00022
g00023	g00023
g00024	g00024
g00025	g00025
g00026	g00026
g00027	g00027
g00028	g00028
g00029	g00029
g00030	g00030
g00031	g00031
g00032	g00032
g00033	g00033
g00034	g00034
g00035	g00035
g00036	g00036
g00037	g00037
g00038	g00038
g00039	g00039
g00040	g00040
g00041	g00041
g00042	g00042
g00043	g00043
g00044	g00044
g00045	g00045
g00046	g00046
g00047	g00047
g00048	g00048
g00049	g00049
g00050	g00050
g00051	g00051
g00052	g00052
g00053	g00053
g00054	g00054
g00055	g00055
g00056	g00056
g00057	g00057
g00058	g00058
g00059	g00059
g00060	g00060
g00061	g00061
g00062	g00062
g00063	g00063
g00064	g00064
g00065	g00065
g00066	g00066
g00067	g00067
g00068	g00068
g00069	g00069
g00070	g00070
g00071	g00071
g00072	g00072
g00073	g00073
g00074	g00074
g00075	g00075
g00076	g00076
g00077	g00077
g00078	g00078
g00079	g00079
g00080	g00080
g00081	g00081
g00082	g00082
g00083	g00083
g00084	g00084
g00085	g00085
g00086	g00086
g00087	g00087
g00088	g00088
g00089	g00089
g00090	g00090
g00091	g00091
g00092	g00092
g00093	g00093
g00094	g00094
g00095	g00095
g00096	g00096
g00097	g00097
g00098	g00098
g00099	g00099
g00100	g00100

ORIGIN

Query Match	24.0%;	Score 215.4;	DB 8;	Length 1085;
Best Local Similarity	73.2%;	Pred. No. 1.8e-42;		
Matches 276: Conservative	0;	Mismatches 101;	Indels 0;	Gaps 0;

Qy	242	CCTGACTACAAATTCCTGATGCTAAATTTCTACAAAGTGGAGCAATTCACAGGCCCTCGCGGA	301
Db	327	CCGATTATGTGCCAGAGGCCAACTTTTACAAAGTAGAAGCAATTTGAGGCCATGCGCGC	386
Qy	302	GTCTCGCAAGTTTCCTCGGCTTTGCTAAATAATGGTATTCGAGGTGTACTGTGTTCTCGAT	361
Db	387	ATCTCCCATGTGACTACGGGTCTATTGGAAATGGGGATTCGTGCGCTAACTGTCTCTGAT	446
Qy	362	GTTTCGAGGTTTTGGTGTCTCAAGGTGGTTCAACTCGAGAGCGAGGCGCTCAGAAATTTTCT	421
Db	447	GTTAGAGGTTTTGGAGTTTCAGGCTGGGATCTGCAGAACGGCAAGCAGGCTCCGAGGTTTTCT	506
Qy	422	GAAGCAAGTTTGTGTCTAAAGTTAAGATCGGAGATCGTGGTTAGCAAGAACCAGGTTGAG	481
Db	507	AAAGACAAATTTGTGTCAAAAGTAAAGATCGGAGTTGTGGTATCTAAAGATCAGGTAGAA	566
Qy	482	GATGTTTATGAAAAAATCATTGGAGGAGGCAAGAACTGGAGAGATTGGAGACGCCAAGATT	541
Db	567	GCAGTAATTGATGCAATTCATTGATGAGGCAAGAACTGGAGAAATTTGGAGATTCGCAAAATA	626
Qy	542	TTCTTGCTGCTGTTTCAGATGTATAAGAGTCCGCACTGGTGAGCGGGGTGTAGAGGCT	601
Db	627	TTTGTTGTTCCAGTTTCAGATGTCATTCGTGTGCAAGACAGTGCAGCGGGACTTTGAAGCCA	686
Qy	602	GAGAGGATGACAGAGG	618
Db	687	GAGAGATGGCTGGTGG	703

RESULT 14
 AF017419 686 bp DNA linear BCT 15-JUL-1998
 LOCUS Nostoc punctiforme PII (glnB) gene, complete cds.
 DEFINITION
 ACCESSION AF017419
 VERSION AF017419.1 GI:2394287
 KEYWORDS
 SOURCE Nostoc punctiforme
 ORGANISM
 Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 REFERENCE
 1 (bases 1 to 686)
 Hanson, T.E., Forchhammer, K., Tandeau de Marsac, N. and Meeks, J.C.
 AUTHORS Characterization of the glnB gene product of Nostoc punctiforme
 TITLE strain ATCC 29133: glnB or the PII protein may be essential
 JOURNAL Microbiology 144 (Pt 6), 1537-1547 (1998)
 MEDLINE 98304077
 PUBMED 9639924
 REFERENCE
 2 (bases 1 to 686)
 Hanson, T.E. and Meeks, J.C.
 AUTHORS Direct Submission
 TITLE Submitted (05-AUG-1997) Section of Microbiology, University of
 JOURNAL California, One Shields Avenue, Davis, CA 95616, USA
 FEATURES
 Location/Qualifiers
 source
 1. .686
 /organism="Nostoc punctiforme"
 /mol_type="genomic DNA"
 /strain="ATCC29133; PCC 73102"
 /db_xref="ATCC:29133"
 /db_xref="taxon:63737"
 275. .613
 /gene="glnB"
 275. .613
 /gene="glnB"
 /function="putative nitrogen regulatory protein"
 /codon_start=1
 /transl_table=11
 /product="PII"
 /protein_id="AAC36348.1"
 /db_xref="GI:2394288"
 /translation="MKKVEAIIRPKLDEVKIALVNAVIGMTVSEVGFGRQKQTE
 RYRSGEYTFEFLQKLKVEIIVDDNQVMVDKIIAAARTGEIGDKIFISPVQVIRI
 RTGKNTAV"

ORIGIN

Query Match 11.9%; Score 107; DB 1; Length 686;
 Best Local Similarity 60.4%; Pred. No. 1.2e-15;
 Matches 195; Conservative 0; Mismatches 125; Indels 3; Gaps 1;
 QY 263 AATTTACAAAGTGAAGCAATCTCAGGCCCTGGCGAGTCTCGCAAGTTTCCTCGGCT 322
 DB 272 AAGATGAAGAGTAGAGCTATTATCCGCCATTAAAGCTAGATGAGGTAAATTCCT 331
 QY 323 TTGCTAAAAATTGGATTCGAGGTGTTACTGTTCTGATGTCGAGGTTTGGTGCTCAA 382
 DB 332 TTGGTTAACGCTGATTTGTCGTATGACTGTTCTGAAGTTCGGGGGTTTCGACGCAA 391
 QY 383 GGTGTTCAACTGAGACGCGGCTCAGAAATTTCTGAACACAGTTTGTCTCTAAA 442
 DB 392 AAGGCCCAACTGAACGGTATCGGGTTCTGAGTACACCGTTG---AGTTTTCGAAAA 448
 QY 443 GTTAAGATGGAGATCGTGGTTAGCAAGACAGGTTGAGGATGTTATAGAAAAATCAAT 502
 DB 449 CTCAAAGTGAAGATCGTGTTCAGCACAATCAGTTGATATCGTGTAGACAAATTTAT 508
 QY 503 GAGGAGCAAGACTGGAGATTTGAGACGCGCAAGATTTTCTGCTGCTGTTTCAGAT 562
 DB 509 GCTGCTGCCCGCACTGGTGAATCGCGATGTAATAATTTTCACTCTGCTGTTGAGCAA 568
 QY 563 GTAATAAGATTCGACATGGTGA 585
 DB 569 GTGATTCGGATTCTGTAATCGGCA 591

RESULT 15
 ASP251822 2493 bp DNA linear BCT 15-FEB-2000
 LOCUS Anabaena sp. PCC7120 glnB gene for PII protein.
 DEFINITION
 ACCESSION AJ251822
 VERSION AJ251822.1 GI:6983586
 KEYWORDS glnB gene; PII protein.
 SOURCE Nostoc sp. PCC 7120 (Anabaena sp. PCC 7120)
 ORGANISM
 Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 REFERENCE
 1
 Gonzalez, L., Phalip, V. and Zhang, C.C.
 AUTHORS Phosphorylation of the signal transduction protein PII by the
 TITLE Ser/Thr kinase PknC in the cyanobacterium Anabaena sp. strain PCC
 7120
 JOURNAL Unpublished
 REFERENCE
 2 (bases 1 to 2493)
 Zhang, C.C.
 AUTHORS Direct Submission
 TITLE Submitted (14-DEC-1999) Zhang C.C., Laboratoire de Chimie
 JOURNAL Bacterienne, C.N.R.S., 31, chemin Joseph Aiguier, 13402 Marseille
 cedex 20, FRANCE
 FEATURES
 Location/Qualifiers
 source
 1. .2493
 /organism="Nostoc sp. PCC 7120"
 /mol_type="genomic DNA"
 /strain="PCC 7120"
 /db_xref="taxon:103690"
 1074. .1412
 /gene="glnB"
 1074. .1412
 /gene="glnB"
 /function="signaling protein"
 /codon_start=1
 /transl_table=11
 /product="PII protein"
 /protein_id="CAB75358.1"
 /db_xref="GI:6983587"
 /db_xref="GOA:Q9L422"
 /db_xref="SPTREMBL:Q9L422"
 /translation="MKKVEAIIRPKLDEVKIALVNAVIGMTVSEVGFGRQKQTE
 RYRSGEYTFEFLQKLKVEIIVDDNQVMVDKIIAAARTGEIGDKIFISPVQVIRI
 RTGKNTAV"

ORIGIN

Query Match 11.5%; Score 103.2; DB 1; Length 2493;
 Best Local Similarity 58.5%; Pred. No. 1.1e-14;
 Matches 199; Conservative 0; Mismatches 138; Indels 3; Gaps 1;
 QY 265 ATCTACAAAGTGAAGCAATCTCAGGCCCTGGCGAGTCTCGCAAGTTTCCTCGGCTTT 324
 DB 1073 AATGAAAAGTAGAGCTATTATCCGCCATTAAAGCTAGATGAGTGAAATCGCTTT 1132
 QY 325 GCTAAAAATTGGTATTCGAGGTGTTACTGTTCTGATGTCGAGTTTGGTGCTCAAG 384
 DB 1133 AGTCAACGCGGTATTGTGGGTATGACTGTTTCTGAAGTTCGGGGTTTGGACGCGAA 1192
 QY 385 TGGTCAACTGAGCGAGCGGCTCAGAAATTTCTGAACACAGTTTGTGCTAAAGT 444
 DB 1193 AGGACAGACAGACGCTATCGCGCTCTGAGTACACTGTG---AGTTTCTGCAAAACT 1249
 QY 445 TAAGATGAGATCGTGGTTAGCAAGACAGGTTGAGGATGTTATAGAAAAATCATTTGA 504
 DB 1250 GAAGTGGAGATTGATTGTAAGATAACCAAGTTGATATGTTGTCACAAATCATTTG 1309
 QY 505 GGAGCAAGAACTGGAGATTGGAGACGCGCAAGATTTTCTGCTGCTGTTTCAGATG 564
 DB 1310 TGCTGCTGTCAGCGCAATCGGTGATGTTGTAATAATTTTATCTCACCTGTAGACAAGT 1369
 QY 565 AATAGAGTCCGACTGTGTAGCGGGGTGATAAGCTGAG 604
 DB 1370 AATTCCGATTCTGTATCTGGAGAGAGAAATACAGAGCCCGTG 1409

Search completed: May 27, 2004, 22:22:42
Job time : 2654.6 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2004, 13:28:59 ; Search time 312.835 Seconds

(without alignments)
12180.972 Million cell updates/sec

Title: us-09-756-541-14

Perfect score: 897

Sequence: 1 GCGGTCTGCGCCCTCTAGA.....AAATAAAATAAACTCGAG 897

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 29Jan04:*

1: geneseqn1980s:*

2: geneseqn1990s:*

3: geneseqn2000s:*

4: geneseqn2001as:*

5: geneseqn2001bs:*

6: geneseqn2002as:*

7: geneseqn2003as:*

8: geneseqn2003bs:*

9: geneseqn2003cs:*

10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	897	100.0	897	4	Aaf58582 Ricinus c
2	594	66.2	594	4	Aaf58584 Ricinus c
3	286.4	31.9	817	4	Aaf58581 Arabidops
4	285	31.8	588	4	Aaf58583 Arabidops
5	285	31.8	591	6	Abz14328 Arabidops
6	96.4	10.7	110000	2	Continuation (2 of
7	96.2	10.7	110000	2	Continuation (15 o
8	84.8	9.5	96109	4	Aaf28548 Genomic f
9	83.6	9.3	375	8	Ada30024 DNA encod
10	81.4	9.1	339	7	Acf72027 Phototrab
11	81.4	9.1	110000	7	Continuation (53 o
12	81.4	9.1	110000	7	Continuation (4 of
13	77.6	8.7	110000	2	Continuation (4 of
14	77	8.6	363	7	Abz40022 N. gonorr
15	73	8.1	16526	3	Aaa81472 N. mening
16	73	8.1	110000	3	Continuation (7 of
17	73	8.1	172325	3	Aaf21613 Neisseria
18	73	8.1	349980	3	Aaf21612 Neisseria
19	70.8	7.9	336	5	Aah68445 C glutami
20	70.8	7.9	417	5	Aah67222 C glutami
21	70.8	7.9	349980	5	Aah68531 C glutami
22	69	7.7	1015	6	Abn74670 Bovine em
23	68.8	7.7	999	6	Abn74679 Bovine em

24	65.2	7.3	770	5	ABA13777 Human ner
25	64.8	7.2	290	4	Aaf71296 Corynebac
26	64.6	7.2	654	4	Aah34361 Human col
27	64.2	7.2	1164	5	Aas88441 DNA encod
28	64.2	7.2	11095	4	Aas46256 DNA encod
29	64	7.1	939	6	Abn74051 Bovine em
30	63.8	7.1	845	5	Aad05422 Human gen
31	63.8	7.1	845	7	Acc50696 Human sec
32	63.8	7.1	845	7	Abz71380 Secreted
33	63.8	7.1	845	8	Adb91328 Human sec
34	63.8	7.1	845	9	Adc73801 Human sec
35	63.6	7.1	958	6	Abn74057 Bovine em
36	63.6	7.1	998	6	Abn74047 Bovine em
37	63.4	7.1	973	6	Abn74320 Bovine em
38	63.2	7.0	404	6	AbL69195 Prostate
39	63.2	7.0	404	6	AbK64407 Human ben
40	62.8	7.0	882	3	Aac63454 Human sec
41	62.8	7.0	882	3	Abz73574 Secreted
42	62.8	7.0	882	7	Ada98087 Human sec
43	62.8	7.0	1012	4	Abz77106 Prolifera
44	62.8	7.0	2163	4	Aaf91900 Human sec
45	62.6	7.0	570	4	Aah33656 Human col

ALIGNMENTS

RESULT 1

AAFS8582
ID AAF58582 standard; cDNA; 897 BP.
XX
AC AAF58582;
XX
DT 23-APR-2001 (first entry)
XX
DE Ricinus communis P-PII cDNA.
XX
KW Castor bean; PII; plant nitrogen regulatory gene; P-PII;
KW nitrogen assimilation; transgenic plant; herbicide screening; ss.
XX
OS Ricinus communis.
XX
PN US6177275-B1.
XX
PD 23-JAN-2001.
XX
PF 23-JUL-1997; 97US-00899330.
XX
PR 24-JUL-1996; 96US-0022328P.
XX
PA (UTNY) UNIV NEW YORK STATE.
XX
PI Coruzzi GM, Lam H, Hsieh M;
XX
DR WPL; 2001-158572/16.
XX
P-PSDB; AAB69496.
XX
PT Novel P-PII genes capable of regulating plant nitrogen assimilation,
PT useful for transgenic plant production, and as probes for isolating
PT additional genomic clones having P-PII gene promoters.
XX
PS Claim 2; Fig 13; 35pp; English.
XX
CC The present sequence encodes a nitrogen regulatory PII protein. Novel
CC plant PII (also called P-PII) nucleotide sequences have been isolated.
CC They are useful for regulating nitrogen assimilation in plants, and in
CC transgenic plant production. They are also used to engineer organisms
CC that overexpress wild-type or mutant P-PII regulatory proteins. P-PII
CC proteins are useful for in vitro screening of herbicides. P-PII
CC nucleotides may be used as probes for isolating additional genomic clones
CC with the promoters of P-PII genes. P-PII promoters are light- and/or
CC sucrose-inducible, and are suitable for genetic engineering of plants

RESULT 2
AAF58584
ID AAF
XX
AC AAF
XX

```

QY 470 GACCAAGTTGAGGATGTTTATAGAAAATCATGAGGAGGCAAGAACTCGAGAGATTGGA 529
DB 421 GACCAAGTTGAGGATGTTTATAGAAAATCATGAGGAGGCAAGAACTCGAGAGATTGGA 480
QY 530 GACGCAAGATTTCTGCTGCTCTTTTCAGATGTAATAAGAGTCCGCACTGAGCGG 589
DB 481 GACGCAAGATTTCTGCTGCTCTTTTCAGATGTAATAAGAGTCCGCACTGAGCGG 540
QY 590 GGTGATAGGCTGAGAGGATGACAGGAGGCGGATCTGACATGAGTACTTCTGCT 643
DB 541 GGTGATAGGCTGAGAGGATGACAGGAGGCGGATCTGACATGAGTACTTCTGCT 594

RESULT 3
AAF58581
ID AAF58581 standard; cDNA; 817 BP.
XX
AC AAF58581;
XX
DT 23-APR-2001 (first entry)
DE Arabidopsis thaliana P-P-II cDNA.
XX
KW Arabidopsis thaliana; P-II; plant nitrogen regulatory gene; P-P-II;
KW nitrogen assimilation; transgenic plant; herbicide screening; ss.
XX
OS Arabidopsis thaliana.
XX
PN US6177275-B1.
XX
PD 23-JAN-2001.
XX
PF 23-JUL-1997; 97US-00899330.
XX
PR 24-JUL-1996; 96US-0022328P.
XX
PA (UUNY ) UNIV NEW YORK STATE.
XX
PI Coruzzi GM, Lam H, Hsieh M;
XX
DR WPI; 2001-158572/16.
DR P-PSDB; AAB69495.
XX
PT Novel P-P-II genes capable of regulating plant nitrogen assimilation,
PT useful for transgenic plant production, and as probes for isolating
PT additional genomic clones having P-P-II gene promoters.
XX
PS Claim 2; Fig 12; 35pp; English.
XX
CC The present sequence encodes a nitrogen regulatory P-II protein. Novel
CC plant P-II (also called P-P-II) nucleotide sequences have been isolated.
CC They are useful for regulating nitrogen assimilation in plants, and in
CC transgenic plant production. They are also used to engineer organisms
CC that overexpress wild-type or mutant P-P-II regulatory proteins. P-P-II
CC proteins are useful for in vitro screening of herbicides. P-P-II
CC nucleotides may be used as probes for isolating additional genomic clones
CC with the promoters of P-P-II genes. P-P-II promoters are light- and/or
CC sucrose-inducible, and are suitable for genetic engineering of plants
XX
SQ Sequence 817 BP; 242 A; 137 C; 188 G; 250 T; 0 U; 0 Other;

Query Match 31.9%; Score 286.4; DB 4; Length 817;
Best Local Similarity 69.8%; Pred. No. 7.7e-61;
Matches 402; Conservative 0; Mismatches 171; Indels 3; Gaps 1;

QY 45 GAATTCGGACGAGGCTACTGCGAACTGGGCTGCTCACTCTCTTCATCTTAATAACA 104
DB 28 GAATTCATGGCGGCTCAATGAAGAAACCCATCTCAATAACTTCTCTGGGTTCTATTCTG 87
QY 105 TCAAGAAAGAAATCCCTGTTTGTGTTTATTCAGTTGTTTTCAGAGCTTAGACATCTC 164
DB 88 ATCGAAGAACAATGCTTTCTCTGATTCGATTTTCGATTTGTTTCGATTCAGACATCCC 147

```

```

QY 165 GGTTCCTCCTTAACACGCGGTCAAGCGGCTAAGATATGCC--CCCGTCGTTCCCTG 221
DB 148 GACCATCTTCCTCGATTTGGTGCACAAAGTCAACGAGTAAATACAGTCGTTGTTTACCTG 207
QY 222 TGATTAATGCCAAAGCTCGCCTGACTACATTCCTCATGCTAAATTTCTACAAAGTGAAG 281
DB 208 TCGTTAGTCCCAATATCTTCGATTAATTCAGACTCGAATTTTACAAGSTGAAG 267
QY 282 CAATTCAGAGCCCTCGCAGTCTCGCAAGTTTCCTCGGCTTTGCTAAATAATTCGTTATTC 341
DB 268 CAATTCAGAGCCATCGCAAGTCCAGCAAGTTTCATCGGCTTTACTGAAATTCGGGATTC 327
QY 342 GAGGTGTTACTGTTCTGATGTCGAGGTTTTCGCTCAAGTGGTTCACACTGAGAGGC 401
DB 328 GAGGTGTTACTGTTCTGATGTCGAGGTTTTCGTCACAGGAGGTTCTTACCGAGAGAC 387
QY 402 AGGGCGGCTCAGAAATTTCTGAAGACAAAGTTTCTGCTAAAGTTAAAGATCGAGATCGTG 461
DB 388 ACGGTGGCTCTGAGTTCTCGAAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTG 447
QY 462 TTAGCAAGACCAAGGTGAGGATGTTTATAGAAAAAATCATTGAGGAGGCAAGAACTCGAG 521
DB 448 TTAAGAAAGACCAAGTGAATCTGTAATCAACAATAATTAAGGAGCAAGGACAGGAG 507
QY 522 AGATTGAGACGCAAGATTTCTTCTGCTGCTGTTTCAGATGTAATAAGAGTCGCACTG 581
DB 508 AGATTGCTGATGGCAAGATTTTGTGTTGCTGCTGTCAGATGTCATAAGAGTTAGGACAG 567
QY 582 GTGAGCGGGCTGATAAGGCTGAGAGGATGACAGGAG 617
DB 568 GTGAGCGGTGGGAGAAAGCAGACAGATGACTGCTG 603

RESULT 4
AAF58583
ID AAF58583 standard; cDNA; 588 BP.
XX
AC AAF58583;
XX
DT 23-APR-2001 (first entry)
DE Arabidopsis thaliana P-P-II cDNA fragment.
XX
KW Arabidopsis thaliana; P-II; plant nitrogen regulatory gene; P-P-II;
KW nitrogen assimilation; transgenic plant; herbicide screening; ss.
XX
OS Arabidopsis thaliana.
XX
PN US6177275-B1.
XX
PD 23-JAN-2001.
XX
PF 23-JUL-1997; 97US-00899330.
XX
PR 24-JUL-1996; 96US-0022328P.
XX
PA (UUNY ) UNIV NEW YORK STATE.
XX
PI Coruzzi GM, Lam H, Hsieh M;
XX
DR WPI; 2001-158572/16.
DR P-PSDB; AAB69495.
XX
PT Novel P-P-II genes capable of regulating plant nitrogen assimilation,
PT useful for transgenic plant production, and as probes for isolating
PT additional genomic clones having P-P-II gene promoters.
XX
PS Claim 2; Col 37-38; 35pp; English.
XX
CC The present sequence encodes a nitrogen regulatory P-II protein. Novel
CC plant P-II (also called P-P-II) nucleotide sequences have been isolated.
CC They are useful for regulating nitrogen assimilation in plants, and in

```


CC specification e.g. is useful for identifying diagnostic and therapeutic
 CC compositions. M. catarrhalis (Branhamella catarrhalis) is a large
 CC aerobic, gram-negative diplococcus, normally found among the bacterial
 CC flora of human upper airways. M. catarrhalis is known to cause acute,
 CC localised infections such as otitis media, sinusitis and bronchopulmonary
 CC infection and life-threatening, systemic diseases including endocarditis
 CC and meningitis
 XX
 XX Sequence 96109 BP; 28783 A; 18910 C; 20341 G; 28075 T; 0 U; 0 Other;

Query Match 9.5%; Score 84.8; DB 4; Length 96109;
 Best Local Similarity 56.2%; Pred. No. 3.1e-12;
 Matches 180; Conservative 0; Mismatches 137; Indels 3; Gaps 1;
 QY 281 GCAATTCACGGCCCTGGCGAGTCTCGCAAGTTTCCTGCGCTTTGCTAAATAATTGGTATT 340
 Db 12953 GCAATTATCAAGCGGTTTAAACTCGATGATGCGGAGAGCACTCTCAGAAATTGGGCTC 13012
 QY 341 CGAGTGTTACTGTTTCGATGTTGCGAGTTTTCGTTGCTCAAGTGGTTCACCTGAGAGG 400
 Db 13013 AATGGTATCACCGTCACTGAGTCAAGGCTTTGGTCGCCAAAGGTCATACCGAGATG 13072
 QY 401 CAGGCGGCTCAGAAATTTCTGGAAGCAAGTTTGTGCTAAAGTTAAGATGGAGATCGTG 460
 Db 13073 TATCGTGGGCGGATATGTTGGTTGAT---TTTTACCAAAATTAAATTGAGATACCA 13129
 QY 461 GTTAGCAAGACCGAGTTGAGGATGTTATAGAAAAATCATGAGGAGGCAAGAACTGGA 520
 Db 13130 TGTCGTGATGAGATGGTTGATTCAATTATGAGTCAATCATTAAGTTGCAATACAGGT 13189
 QY 521 GAGATTGAGACGCGAAGATTTCTGCTGCTGTTTCAGATGTAATAAGATCGGCACT 580
 Db 13190 AAAATTGGTGTATGATTAAGATTTTGTAGTCCGTTGAGCGGTGTCATTCGCAATCGAAT 13249
 QY 581 GGTGAGCGGGGTGATAAGCC 600
 Db 13250 GCGCAATTGATGAAGTGC 13269

RESULT 9
 ADA30024
 ID ADA30024 standard; DNA; 375 BP.
 XX ADA30024;
 AC
 XX
 XX
 XX 20-NOV-2003 (first entry)
 XX
 XX DNA encoding Acinetobacter baumannii protein #1311.
 XX
 XX ds; gene; Acinetobacter baumannii; bacterial disease; antibacterial;
 XX vaccine; plant biocontrol agent.
 XX
 XX Acinetobacter baumannii.
 XX
 XX
 XX US6562958-B1.
 XX
 XX 13-MAY-2003.
 XX
 XX 04-JUN-1999; 99US-00328352.
 XX
 XX 09-JUN-1998; 98US-0083701P.
 XX
 XX (GENO-) GENOME THERAPEUTICS CORP.
 XX
 XX Breton G, Bush D;
 XX
 XX MPI; 2003-576092/54.
 XX
 XX P-PSDB; ADA34150.
 XX

New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 PT for diagnosing a bacterial disease, as components of antibacterial
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 PT plants.

XX Example; SEQ ID NO 1311; 328pp; English.
 PS
 CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents
 CC for diagnosing a bacterial disease, as components of antibacterial
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 CC A. baumannii and other Acinetobacter species in a sample, in screening
 CC compounds for the ability to interfere with the A. baumannii life cycle
 CC or to inhibit A. baumannii infection, and as biocontrol agents for
 CC plants. The present sequence represents DNA encoding an A. baumannii
 XX protein.
 XX

Sequence 375 BP; 122 A; 64 C; 93 G; 96 T; 0 U; 0 Other;
 Query Match 9.3%; Score 83.6; DB 8; Length 375;
 Best Local Similarity 55.9%; Pred. No. 9.2e-11;
 Matches 180; Conservative 0; Mismatches 139; Indels 3; Gaps 1;
 QY 281 GCAATTCACGGCCCTGGCGAGTCTCGCAAGTTTCCTGCGCTTTGCTAAATAATTGGTATT 340
 Db 52 GCAATTGTAAACCGTTTAAATTGGATGATGTCGCTGAAGCACTCTCTGACATTGGTGA 111
 QY 341 CGAGTGTTACTGTTTCGATGTTTCGAGTTTTCGTTGCTCAAGTGGTTCACCTGAGAGG 400
 Db 112 CAGGGATTACCGTAACTGAAGTTAAAGCTTTTGGTCGTCAAAAGGACATACAGAACTT 171
 QY 401 CAGGCGGCTCAGAAATTTCTGGAAGCAAGTTTGTGCTAAAGTTAAGATGGAGATCGTG 460
 Db 172 TACCGCGCGCTGAGTATGTTGGTTGAT---TCTTACCTAAAGTAAATCGAATTGCG 228
 QY 461 GTTAGCAAGACCGAGTTGAGGATGTTATAGAAAAATCATTGAAGGCAAGAACTGGA 520
 Db 229 ATTAGTGTGAATGGTTCGCGCGGTAATTGAGTCAATTACACGTGTGGCAAGCACTGGA 288
 QY 521 GAGATTGAGACGCGAAGATTTCTGCTGCTGCTTTTCAGATGTAATAAGATCGGCACT 580
 Db 289 AAAATCGCGACGCTAAGATTTTGTGACTAATCTGGAACCAAGTCATCCGTATCCGTACA 348
 QY 581 GGTGAGCGGGGTGATAAGCC 602
 Db 349 GGTGAACACAGGACCGAGATGCTG 370

RESULT 10
 ACF72027
 ID ACF72027 standard; DNA; 339 BP.
 XX ACF72027;
 AC
 XX
 XX 20-NOV-2003 (first entry)
 XX
 XX Photorhabdus luminescens nucleotide sequence #10494.
 XX
 XX Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 XX detection; food; gene expression; plant; animal; microorganism; toxin;
 XX antibiotic; biopesticide; virulence factor; disease model; plague;
 XX whooping cough; gene; ds.
 XX
 XX Photorhabdus luminescens.
 XX
 XX WO200294867-A2.
 XX
 XX 28-NOV-2002.
 XX
 XX 07-FEB-2002; 2002WO-IB003040.
 XX
 XX 07-FEB-2001; 2001PR-00001659.
 XX
 XX (INSP) INST PASTEUR.
 XX PA (CNRS) CNRS CENT NAT RECH SCI.
 XX
 XX Duchaud B, Tacurit S, Glaser P, Frangeul L, Kunst F, Danchin A;
 PI

PI Buchrieser C;
 XX WPI; 2003-148459/14.
 DR
 XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
 XX
 XX Claim 2; SEQ ID NO :0494; 1205pp; French.
 XX
 CC The invention relates to the isolation of genes and their encoded
 CC proteins from Photorhabdus luminescens. The isolated sequences are
 CC sources of probes and primers for detecting the genome of P. luminescens
 CC and related species; to study polymorphisms; for gene analysis and for
 CC detection/amplification of the genes. Antibodies (Ab) raised against the
 CC polypeptides encoded by the genes are used for detection/identification
 CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 CC carry a gene-containing vector are used to select compounds that
 CC modulate, regulate, induce or inhibit expression of the genes in plants,
 CC animals or microorganisms other than P. luminescens and are able to alter
 CC response or sensitivity to toxins and antibiotics produced by P.
 CC luminescens. Cells transformed to express the genes are useful for
 CC recombinant production of the proteins, particularly toxins and
 CC antibacterials useful as insecticides, bactericides and fungicides. The
 CC genes, proteins, vectors containing the genes and Ab are also useful
 CC therapeutically (to treat microbial infection by bacteria or fungi that
 CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
 CC biopesticides. Other uses of the genes and the proteins are as virulence
 CC factors and for identifying targets of human diseases for which P.
 CC luminescens is a model (particularly plague and whooping cough). This
 CC sequence represents one of the isolated P. luminescens genes
 XX
 SQ Sequence 339 BP; 105 A; 51 C; 96 G; 87 T; 0 U; 0 Other;
 Query Match 9.1%; Score 81.4; DB 7; Length 339;
 Best Local Similarity 55.0%; Pred. No. 3.1e-10;
 Matches 182; Conservative 0; Mismatches 146; Indels 3; Gaps 1;
 QY 270 ACAAGTGGAGCAATCTCAGGCGCTCGGAGTCTCGAGTTCCTCGGCTTTGCTAA 329
 Db 5 AAAAGATTGATCGGATTATCAAACTTCAATAGATGATGCGTGAAGCTCTGCGG 64
 QY 330 AAATTGCTATTCGAGGTGTACTGTTCTGATGTTGCGAGTTTGTGCTCAAGTGGTT 389
 Db 65 AAGTGGTATCCCGAATGACGATACAGAGTGAAGTTTGGCGCCAAAGGTC 124
 QY 390 CAAGTGGAGCGGCGGCTGAGATTCTTCTGAGCAAGATTGTTGCTTAAAGTTAAGA 449
 Db 125 ATACAGAGCTGATCGCGGTGACGAATATATGTTGGAT---TTCTGCCAAAGTGAAA 181
 QY 450 TGGAGATCGTGTAGCAAGACCGGTTGAGATGTTATAGAAAATCATTTGAGGAGG 509
 Db 182 TAGAAATTGCTGCGCAGATGATTTGTCGATACCTGTTGTAACCAITATGACAGCG 241
 QY 510 CAAGAACTGGAGATGAGAGCGGCAAGATTCTTCTGCTGCTTTCAGATGTAATA 569
 Db 242 CACAGACCGGAAATCGGTGATGTAAATATTTGTTATTTGATGATGACGTTGTC 301
 QY 570 GAGTGGCGATCGTGTAGCGGGGTGATAAGC 600
 Db 302 GTATCGCACCGGTGAGCAGGATGAAGAGC 332
 RESULT 11
 ACF67367_52
 Continuation (53 of 57) of ACF67367 from base 5200001 (Photorhabdus luminescens nucleoti
 WP Sequence split into 57 fragments LOCUS ACF67367 Accession Acf67367
 WP Fragment Name Begin End
 WP ACF67367_00 1 110000
 WP ACF67367_01 100001 210000
 WP ACF67367_02 200001 310000
 WP ACF67367_03 300001 410000
 WP ACF67367_04 400001 510000
 WP ACF67367_05 500001 610000

WP ACF67367_06 600001 710000
 WP ACF67367_07 700001 810000
 WP ACF67367_08 800001 910000
 WP ACF67367_09 900001 1010000
 WP ACF67367_10 100001 1110000
 WP ACF67367_11 1100001 1210000
 WP ACF67367_12 1200001 1310000
 WP ACF67367_13 1300001 1410000
 WP ACF67367_14 1400001 1510000
 WP ACF67367_15 1500001 1610000
 WP ACF67367_16 1600001 1710000
 WP ACF67367_17 1700001 1810000
 WP ACF67367_18 1800001 1910000
 WP ACF67367_19 1900001 2010000
 WP ACF67367_20 2000001 2110000
 WP ACF67367_21 2100001 2210000
 WP ACF67367_22 2200001 2310000
 WP ACF67367_23 2300001 2410000
 WP ACF67367_24 2400001 2510000
 WP ACF67367_25 2500001 2610000
 WP ACF67367_26 2600001 2710000
 WP ACF67367_27 2700001 2810000
 WP ACF67367_28 2800001 2910000
 WP ACF67367_29 2900001 3010000
 WP ACF67367_30 3000001 3110000
 WP ACF67367_31 3100001 3210000
 WP ACF67367_32 3200001 3310000
 WP ACF67367_33 3300001 3410000
 WP ACF67367_34 3400001 3510000
 WP ACF67367_35 3500001 3610000
 WP ACF67367_36 3600001 3710000
 WP ACF67367_37 3700001 3810000
 WP ACF67367_38 3800001 3910000
 WP ACF67367_39 3900001 4010000
 WP ACF67367_40 4000001 4110000
 WP ACF67367_41 4100001 4210000
 WP ACF67367_42 4200001 4310000
 WP ACF67367_43 4300001 4410000
 WP ACF67367_44 4400001 4510000
 WP ACF67367_45 4500001 4610000
 WP ACF67367_46 4600001 4710000
 WP ACF67367_47 4700001 4810000
 WP ACF67367_48 4800001 4910000
 WP ACF67367_49 4900001 5010000
 WP ACF67367_50 5000001 5110000
 WP ACF67367_51 5100001 5210000
 WP ACF67367_52 5200001 5310000
 WP ACF67367_53 5300001 5410000
 WP ACF67367_54 5400001 5510000
 WP ACF67367_55 5500001 5610000
 WP ACF67367_56 5600001 5648894
 Query Match 9.1%; Score 81.4; DB 7; Length 110000;
 Best Local Similarity 55.0%; Pred. No. 2.3e-09;
 Matches 182; Conservative 0; Mismatches 146; Indels 3; Gaps 1;
 QY 270 ACAAGTGGAGCAATCTCAGGCGCTCGGAGTCTCGAGTTCCTCGGCTTTGCTAA 329
 Db 49307 AAAAGATTGATCGGATTATCAAACTTCAATAGATGATGCGTGAAGCTCTGCGG 49366
 QY 330 AAATTGCTATTCGAGGTGTACTGTTCTGATGTTGCGAGTTTGTGCTCAAGTGGTT 389
 Db 49367 AAGTGGTATCCCGAATGACGATACAGAGTGAAGTTTGGCGGCCAAAGGTC 49426
 QY 390 CAAGTGGAGCGGCGGCTCAGAAATTTCTCAAGCAAGTTTGTGCTTAAAGTTAAGA 449
 Db 49427 ATACAGAGCTGTATCGCGGTGAGATATATGTTGGAT---TTCTGCCAAAGTGAAA 49483
 QY 450 TGGAGATCGTGTAGCAAGACCGGTTGAGATGTTATAGAAAATCATTTGAGGAGG 509
 Db 49484 TAGAAATTGCTGCGCAGATGATTTGTCGATACCTGTTGTAACCAITATGACAGCG 49543
 QY 510 CAAGAACTGGAGATGAGAGCGGCAAGATTCTTCTGCTGCTTTCAGATGTAATA 569

Db 49544 CACAGACCGGGAATCGGTGATGTAAATATTTGTTATGATGACGCTGTGTGTC 49603
QY 570 GAGTCCGCACTGGTGGAGCGGGTGATAAGGC 600
Db 49604 GTATCCGCAACCGGTGAGCAGGATGAAGAGGC 49634

RESULT 12

ACF65387_3
Continuation (4 of 7) of ACF65387 from base 300001 (Photorhabdus luminescens nucleotide
WP Sequence split into 7 fragments LOCUS ACF65387 Accession ACF65387

WP Fragment Name Begin End
WP ACF65387_0 1 110000
WP ACF65387_1 100001 210000
WP ACF65387_2 200001 310000
WP ACF65387_3 300001 410000
WP ACF65387_4 400001 510000
WP ACF65387_5 500001 610000
WP ACF65387_6 600001 696798

Query Match 9.1%; Score 81.4; DB 7; Length 110000;
Best Local Similarity 55.0%; Pred. No. 2.3e-09;
Matches 182; Conservative 0; Mismatches 146; Indels 3; Gaps 1;

QY 270 ACAAGTGGAGCAATCTCAGGCCCTGCGAGTCTCGCAAGTTTCCTCGGCTTTCCTAA 329
Db 88821 AAAAGATTGATGCGATTATCAACCTTTCAAATTAGATGATGCGTGAAGCTCTGGCGG 88880
QY 330 AAATTGGTATCGAGCTGTACTGTCTTCTGATGTTTCGAGTTTTCGTCTCAAGTGGTT 389
Db 88881 AAGTGGGTATCCCGGATGACAGTAAACAGAGTGAAGGTTTGGCGCCAAAAGGTC 88940
QY 390 CAACTGAGAGGCGGCGGCTCAGAAATTTCTGAAGACAAGTTTGTGCTAAAGTTAAGA 449
Db 88941 ATACAGAGCTGTATCGCGTGCAGAAATATATGCTGGAT---TTTCTGCCAAAAGTGA 88997
QY 450 TGGAGATCGTGTAGCAAGACCAAGCTGAGATGTTATAGAAAATCATCTGAGGAGG 509
Db 88998 TAGAAATTTGTGTCGCCAGATGATATGTGCTGATCTGTGTTTCAAGACCATATG 89057
QY 510 CAAGAACTGGAGAGATGGAGACGCGCAGAGATTTCTGTGCTGCTTTTCAGATGTAATAA 569
Db 89058 CACAGACCGGGAATCGGTGATGTAAATATTTGTTATGATGATGACGCTGTGTCG 89117
QY 570 GAGTCCGCACTGGTGGAGCGGGTGATAAGGC 600
Db 89118 GTATCCGCAACCGGTGAGCAGGATGAAGAGGC 89148

RESULT 13

AAT42063_03
Continuation (4 of 19) of AAT42063 from base 300001 (Haemophilus influenzae complete gen
WP Sequence split into 19 fragments LOCUS AAT42063 Accession Aat42063

WP Fragment Name Begin End
WP AAT42063_00 1 110000
WP AAT42063_01 100001 210000
WP AAT42063_02 200001 310000
WP AAT42063_03 300001 410000
WP AAT42063_04 400001 510000
WP AAT42063_05 500001 610000
WP AAT42063_06 600001 710000
WP AAT42063_07 700001 810000
WP AAT42063_08 800001 910000
WP AAT42063_09 900001 1010000
WP AAT42063_10 1000001 1110000
WP AAT42063_11 1100001 1210000
WP AAT42063_12 1200001 1310000
WP AAT42063_13 1300001 1410000
WP AAT42063_14 1400001 1510000
WP AAT42063_15 1500001 1610000
WP AAT42063_16 1600001 1710000
WP AAT42063_17 1700001 1810000

WP AAT42063_18 1800001 1830121
Query Match 8.7%; Score 77.6; DB 2; Length 110000;
Best Local Similarity 55.1%; Pred. No. 2e-08;
Matches 174; Conservative 0; Mismatches 139; Indels 3; Gaps 1;

QY 270 ACAAGTGGAGCAATCTCAGGCCCTGCGAGTCTCGCAAGTTTCCTCGGCTTTCCTAA 329
Db 63919 AAAAATCGAAGCAATGATTAACCTTTAAATATAGACGATGCGAGAAAGTCITTCAG 63978
QY 330 AAATTGGTATTCGAGTGTACTGTCTTCTGATGTTTCGAGGTTTTCGTGCTCAAGGTGGTT 389
Db 63979 ATATTGGTATTTTCAGSTATGACAATCACAGAAAGTACGGGATTTGCTCGTCAAAAAGGTC 64038
QY 390 CAACTGAGAGGCGGCGGCTCAGAAATTTCTGAAGACAAGTTTGTGCTAAAGTTAAGA 449
Db 64039 ATACAGAACTTTATCGTGGTGGCGAATATATGCTGGAT---TTTCTGCCAAAAGTGA 64095
QY 450 TGGAGATCGTGTAGCAAGACCAAGCTTGAAGATGTTATAGAAAATCATCTGAGGAGG 509
Db 64096 TGAAGTGGTATCTCTGATGAGCTTGTGGATCAATGATTAAGCGATATTTGAACGG 64155
QY 510 CAAGAACTGGAGAGATGGAGACGCGCAGAGATTTCTGTGCTGCTGTTTCAGATGTAATAA 569
Db 64156 CACAAACAGGTAATCGGTGACGCGCAAAATTTTGTGTTATCACGTTGAGAGAGCCATCC 64215
QY 570 GAGTCCGCACTGGTGA 585
Db 64216 GCATTCGCAACGGCGA 64231

RESULT 14

ABZ40022
ID ABZ40022 standard; DNA; 363 BP.

XX AC ABZ40022;
XX DT 07-MAR-2003 (first entry)
XX DE N. gonorrhoeae nucleotide sequence SEQ ID 4633.
XX KW Antibacterial; infection; vaccine; gene therapy; gene; ds.
XX OS Neisseria gonorrhoeae.
XX PN WO200279243-A2.
XX PD 10-OCT-2002.
XX PF 12-FEB-2002; 2002WO-IB002069.
XX PR 12-FEB-2001; 2001GB-00003424.
XX PA (CHIR-) CHIRON SPA.
XX PI Fontana MR, Pizza M, Massignani V, Monaci E;
XX WPI; 2003-058415/05.
XX DR P-PSDB; ABP79052.

PT New protein from Neisseria gonorrhoeae, useful for the manufacture of a
PT medicament for treating or preventing N. gonorrhoeae infection.

PS Disclosure; Page 528; 815pp; English.

CC The present invention relates to proteins from Neisseria gonorrhoeae.
CC Also disclosed are the nucleic acid molecules encoding the proteins and
CC antibodies that specifically bind to the proteins. The composition
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing N. gonorrhoeae
CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records ABZ37706-ABZ42016 represent nucleic acid
CC molecules of the invention

```

XX SQ Sequence 363 BP; 84 A; 77 C; 123 G; 79 T; 0 U; 0 Other;
Query Match 8.6%; Score 77; DB 7; Length 363;
Best Local Similarity 54.1%; Pred. No. 3.9e-09;
Matches 180; Conservative 0; Mismatches 150; Indels 3; Gaps 1;

QY 270 ACAAGTGAAGCAATTCAGGCGCTGCGAGTTCGCAAGTTCTCGGCTTGCTAA 329
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
32 AAAAAATCGAGCGATTGTCAAAACGTTCAAGCTGACGACGTGCGCGAGGCGTTGACGG 91
QY 330 AAATGCTATTTCAGGCTTACTGTTCTGANGTTGAGGTTTGTGCTCAAGGTGTT 389
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 92 AAATCGGCATTACGGGCATGACCGTCAGCGAGGTCAAGGTTGCGCAGGCGAGGGGC 151
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 390 CAACTGACAGCGAGGCGCGCTCAGAAATTTCTGAAGACAAAGTTTGTGCTAAAGTTAAGA 449
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
152 ATACGGAAATCTATCGCGTGGGGAATACGCGTCGAT---TTCCTGCCCAAGGTCAAAA 208
QY 450 TGGAGATCGTGTGTAGCAAGACCGAGTTGAGGATGTTATAGAAAAATCATTTGAGGAGG 509
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
209 TCGAGTTGTTGTGGCGGATGATCGGTGGAACGCGCGATTGACGATGATTGTCGAGGTGG 268
QY 510 CAAAGACTGAGAGATTGAGACGCGCAAGATTTCTTGCTGCTGCTGTTTCAGATGTAATAA 569
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
269 CCGGTTGCGGCAAAATCGGCGGACGCAAGATTTTGTGCTGCGCGTCGAGGAGGCAATCC 328
QY 570 GAGTCCGCACTGCTGAGCGGGGTGATAAGGCTG 602
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
329 GTATCGCGACGGCGGACGTTTCGACGCGGCGG 361

RESULT 15
AAAB1472
ID AAB1472 standard; DNA; 16526 BP.
XX AC AAB1472;
XX
XX 04-DEC-2000 (first entry)
XX
XX N. meningitidis partial DNA sequence gnm_20 SEQ ID NO:20.
XX
XX Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
XX antigen; vaccine; diagnosis; infection; antibacterial; identification;
XX Meningococcus B; MenB; ds.
XX
XX Neisseria meningitidis.
XX
XX WO200022430-A2.
XX
XX 20-APR-2000.
XX
XX 08-OCT-1999; 99WO-US023573.
XX
XX 09-OCT-1998; 98US-0103794P.
XX
XX 30-APR-1999; 99US-0132068P.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
XX Massignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scariato V;
XX Rappuoli R, Pizza M;
XX
XX WPI; 2000-318079/27.
XX
XX Isolated nucleotide sequences of Neisseria meningitidis which can be used
XX in the diagnosis and treatment of N. meningitidis infection and other
XX Neisserial infections, for example, N.gonorrhoea.
XX
XX Claim 7; Page 466-471; 1760pp; English.
XX
XX The present invention describes methods of obtaining immunogenic proteins
XX from Neisseria genomic sequences. AAB1453 to AAB1454 represent

```

```

CC specifically claimed Neisseria meningitidis genomic DNA sequences;
CC AAB1260 to AAB1303 and AAB25620 to AAB25663 represent Neisseria DNA
CC sequences and their corresponding proteins; AAB1254 to AAB1259 and
CC AAB1304 to AAB1321 represent PCR primers used in the isolation of
CC Neisseria meningitidis DNA sequences; and AAB1322 to AAB1452 represent
CC Neisseria meningitidis MenB polynucleotide ORF sequences, which are all
CC used in the exemplification of the present invention. The nucleic acid
CC sequences, protein sequences, and antibodies against them, can be used in
CC the manufacture of a composition. The composition can be used as a
CC medicament (or in the manufacture of a medicament) for treating,
CC preventing or diagnosing infection due to Neisserial bacteria. For
CC example, some of the identified proteins could be components of vaccines
CC against Meningococcus B; against all serotypes; and/or against all
CC pathogenic Neisseriae. Identification of sequences from the bacterium
CC will also facilitate production of biological probes, particularly
CC organism-specific probes. Attempts to make efficacious Meningococcus B
CC vaccines have failed mainly due to antigen tolerance. Multivalent
CC vaccines have also been tried but none have successfully overcome
CC antigenic variability. The provision of further, complete sequenced proteins
CC provide an opportunity to identify secreted or surface exposed proteins
CC that may be presumed targets for the immune system and which are not
CC antigenically variable or at least more conserved than other more
CC variable regions
XX
SQ Sequence 16526 BP; 3183 A; 4006 C; 5067 G; 4268 T; 0 U; 2 Other;

Query Match 8.1%; Score 73; DB 3; Length 16526;
Best Local Similarity 55.3%; Pred. No. 1.4e-07;
Matches 184; Conservative 0; Mismatches 145; Indels 4; Gaps 2;

QY 270 ACAAGTGAAGCAATTCAGGCGCTGCGAGTTCGCAAGTTTCTCGGCTTGCTAA 329
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
13603 AAAAAATCGAGCGATTGTCAAAACGTTCAACTGACGACGTGCGCGAGCGTTGACGG 13662
QY 330 AAATGCTATTTCAGGCTTACTGTTTCTGAAGTTTCAGGTTTGTGTCCTCAAGGTGTT 389
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
13663 AAATCGGCATTACGGGCATGACCGTCAGCGAGGTCAAAGGTTTCGCGAGCAGAGGCGC 13722
QY 390 CAACTGAGAGGCGAGGCGGCTCAGAAATTTCTGAAGACAAAGTTTGTGCTAAAGTTAAGA 449
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
13723 ATACGGAAATCTATCGCGCGCGGAAATACCGCTCGAT---TTCCTGCCCAAAATCAAAA 13779
QY 450 TGGAGATCGTGTGTAGCAAGACCAAGTTTCAGGATGTTATAGAAAAATCATTTGAGGAGG 509
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
13780 TCGAGCTGGTGTGCGGATGATCTGTGGAACGCGCGATTGACGTGATTTCGAGGTGG 13838
QY 510 CAAAGAACTGGAGAGATTGGAGACGCGCAAGATTTTTCGTCGCTGTTTCAGATGTAATAA 569
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
13839 CGCGTTTCGGGAAAAATCGCGGACGCGCAAGATTTTGTGCTGCGGTTGAGGAGCAATCC 13898
QY 570 GAGTCCGCACTGCTGAGCGGGGTGATAAGGCTG 602
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
13899 GTATCGCGACGGCGAACGTTTCGAGCGCGCAG 13931

Search completed: May 27, 2004, 19:59:40
Job time : 315.835 secs

```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2004, 19:07:55 ; Search time 2078.03 Seconds
(without alignments)
12890.278 Million cell updates/sec

Title: US-09-756-541-14

Perfect score: 897

Sequence: 1 GCGGTCTGGCGCTCTAGA.....AAAAAAAAAAAAAATCGAG 897

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_hic.*

9: gb_est1.*

10: gb_est2.*

11: gb_hic.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_nam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rtd.*

26: em_gss_pbg.*

27: em_gss_vrt.*

28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	408.8	45.6	542	14	T15252
2	321.4	35.8	782	14	CB289081
3	300.8	33.5	723	13	BO798001
4	298.6	33.3	597	10	BE239550

5	298.6	33.3	714	12	BI311072
6	287	33.1	812	14	CA922829
7	293	32.7	612	10	BF645522
8	284.2	31.7	759	14	CF920467
9	283	31.5	866	13	CB1819
10	277.8	31.0	677	10	BB823431
11	276	30.8	799	12	BM113473
12	275.8	30.7	711	12	BI932336
13	275	30.7	732	9	AJ558383
14	274.2	30.6	613	9	AM035791
15	274.2	30.6	648	9	AI773079
16	274.2	30.6	702	12	BI932123
17	274.2	30.6	741	10	AW738071
18	274.2	30.6	804	12	BI930336
19	274.2	30.6	810	12	BI930060
20	273.8	30.5	756	12	BS959705
21	273.6	30.5	736	12	BI932913
22	273.2	30.5	685	12	BI929601
23	272.6	30.4	630	10	AW929534
24	267	29.8	719	12	BJ573686
25	266.2	29.7	732	14	CD820903
26	265.4	29.6	657	14	CB349285
27	265.4	29.6	701	14	CB349370
28	264.6	29.5	458	10	BE210153
29	264.6	29.5	669	14	CD822497
30	263	29.3	685	14	CD825321
31	263	29.3	745	14	CD819046
32	262.2	29.2	675	14	CB346213
33	249.2	27.8	549	14	CB261070
34	248.2	27.7	560	14	CA514765
35	246.4	27.5	579	10	AW223643
36	244.4	27.2	673	14	CB085731
37	239.2	26.7	540	14	CB349912
38	238.6	26.6	551	14	CB349990
39	238.6	26.6	598	14	CB350294
40	236.4	26.4	540	10	AW624562
41	233.4	26.0	532	14	CB350222
42	227.6	25.4	564	13	BQ118918
43	227.4	25.4	729	14	CF475242
44	227.2	25.3	796	14	CF395143
45	223.6	24.9	690	14	CA290108

ALIGNMENTS

RESULT 1
T15252
LOCUS
DEFINITION
T15252
T15252
T15252.1 GI:14190796
542 bp mRNA linear EST 23-MAY-2001
crr852 lambdaZAPST Ricinus communis cDNA clone pcrs852 similar to
nitrogen-regulatory protein, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Ricinus communis (castor bean)
Ricinus communis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Malpighiales; Euphorbiaceae; Acalyphoideae;
Acalyphaeae; Ricinus.
REFERENCE
1 (bases 1 to 542)
vandeLoe,F.J., Turner,S. and Somerville,C.
Expressed sequence tags from developing castor seeds
Plant Physiol. 108, 1141-1150 (1995)
CONTACT: Somerville CR
Carnegie Institution
Carnegie Institution, 290 Panama St., Stanford, CA 94305
Tel: 4153251521
Email: crs@andrew.stanford.edu
Seq primer: T3.
Location/Qualifiers
1..542
/organism="Ricinus communis"

```

/mol_type="mRNA"
/strain="Baker 296"
/db_xref="taxon:3988"
/clone="pcrs852"
/clone_lib="lambdaZAPST"
/note="vector: lambdaZAPII; Site_1: EcoRI; Site_2: XhoI;
Poly(A) + RNA was purified from developing stage III to
stage V (Greenwood & Bewley, Can. J. Bot. 60:1751-1760,
1992) endosperm plus embryo of immature castor fruits.
cDNA was synthesized and cloned into lambdaZAPII according
to the instructions of the manufacturer (Stratagene).
synthesis was primed from the poly(A) tail, and cloned
directionally into XhoI (3') and EcoRI (5') sites. In few
cases, sequence data indicated that this directionality
was reversed. Partial cDNA clones predominate."

```

ORIGIN

Query Match	45.6%	Score 408.8	DB 14	Length 542
Best Local Similarity	95.3%	Pred. No. 1.7e-54		
Matches 428	Conservative 0	Mismatches 18	Indels 2	Gaps 1
QY	59	GCCTACTCGGAAACTGGCGCTTGCTCACTCCCTCTTCATTCTAATCAATCAAGAAAGAAATTC	118	
Db	1	GCCTACTCGGAAACTGGCGCTTGCTCACTCCCTCTTCATTCTAATCAATCAAGAAAGAAATTC	60	
QY	119	CCCTGTTTTCAGTTTCAGTTTGTGTTGTCAGAGCTTAGACAATTCCTCGTTCCTCAGTTT	178	
Db	61	CCCTGTTTTCAGTTTTCAGTTTGTGTTGTCAGAGCTTAGACAATTCCTCGTTCCTCAGTTT	120	
QY	179	AACACCGCGCTCAAGCGCGTAAGATATGCCCGCTGCTTCCTGTGATTAATGCCCAAAGC	238	
Db	121	AACACCGCGCTCAAGCGCGTAAGATATGCCCGCTGCTTCCTGTGATTAATGCCCAAAGC	180	
QY	239	TCGCCTGACTACATTCCTGTGATGCTAAATCTCAAACTGGAAGCAATTCAGGCCCTCG	298	
Db	181	TCGCCTGACTACATTCCTGTGATGCTAAATCTCAAACTGGAAGCAATTCAGGCCCTCG	240	
QY	299	CGAGTCTCGCAAGTTTCTCGGCTTTGCTCTAAATTCGTATTCGAGTGTTACTGTGTTCT	358	
Db	241	CGAGTCTCGCAAGTTTCTCGGCTTTGCTCTAAATTCGTATTCGAGTGTTACTGTGTTCT	300	
QY	359	GATGTTTCGAGGTTTTGCTGCTCAAGTGGTTTCAACTGAGAGCGCAGCGGCTCAGAAATTT	418	
Db	301	GATGTTTCGAGGTTTTGCTGCTCAAGTGGTTTCAACTGAGAGCGCAGCGGCTCAGAAATTT	360	
QY	419	TCTGAAGACAAGTTTGTCTCTAAGTTTAAGATCGAGATCGTGGTTTTCGAAGACCCAGGTT	478	
Db	361	NCYGAAGNCAGTTTGTGCTTAAAGTAAGATGAGACCGTGGT-NGCAAGCCAGGTT	418	
QY	479	GAGGATCTTATAGAAAAAATCATTGAGG	506	
Db	419	GAGTCTGTTATGCAAAATNATTGGGG	446	

[illegible]

Department of Fruit Science
Southwest Missouri State University-Mountain Grove
9740 Red Spring Road, Mountain Grove, MO 65711-2999, USA
Tel: 417 926 4105
Fax: 417 926 6646
Email: weq070@smsu.edu
Insert Length: 782 Std Error: 0.00
Plate: VAN-Baker-1-12 row: F column: 06
Seq primer: T3 PRIMER
High quality sequence stop: 782
POLYA=No.

FEATURES

source

Location/Qualifiers

```

1. 782
/organism="Vitis aestivalis"
/mol_type="cDNA"
/cultivar="Norton"
/db_xref="taxon:3605"
/clone="V-B-112F06"
/tissue_type="Leaf"
/dev_stage="Young leaf"
/lab_host="XL10-Gold E.coli"
/clone_lib="VAN-Baker-1"
/notes="vector: pBluescript I
I; Site_2: EcoR I; VAN-Baker-1:
grape young leaves (Vitis a
grapevines were grown under
The cDNA synthesis and library
according to the instructions
cDNA library construction kit

```

ORIGIN

Query Match	35.8%	Score	321.4	DB 14	Length	782			
Best Local Similarity	73.7%	Pred. No.	5.7e-43						
Matches	409	Conservative	0	Mismatches	146	Indels	0	Gaps	0

QY	111	AAGAAATCCCGTTTGTGATTTTCAGTTTGTGTTTGTCCAGAGCTTAGACATCTCGGTTT	170
DB	65	AAGAAATGGCTCTCATGAGCTGCACCTTGGATCGGCCCAATCTCAAGAATCTCGAAATT	124
QY	171	CTCACTTTAACCGGGTCAAGGGGTAAAGATATGCCCGCTCGTTCCCTGTGATTAATG	230
DB	125	TCCAGGTTTAATTTGTGCTTAAACGGCTCAAGAATGGATCTGTTCTTCCCATGTCGAG	184
QY	231	CCCAAGAGCTCGGCTGACTACATTTCTTGATGCTAAATTTCTACAAAGTGGAGCAATTTCTCA	290
DB	185	CTCAGAGCTCTCCAGATTATACCCAGACGGCGAGTTTTCACAAAGTTGAAGCGATCTCTGA	244
QY	291	GGCCCTGGCGAGTCTGCGAAGTTTCTTCGAGCTTGGCTAATAAATGGTATTTCGAGTGTTA	350
DB	245	GGCCCTGGCGAATCCAGCAGGTTTCTTTCGGCTTTCGCTGAAATGGGTATTTCGTGGTGT	304
QY	351	CTGTTTCTGATGTTTCGAGGTTTGGTGCTCAAGGTGGTTCAACTGAGAGCGAGGCGGCT	410
DB	305	CTGTTTCTGATGTCCGAGGCTTTGGTGCTCAAGGCGGATCGCCTGAAAGACAGGCAGGCT	364
QY	411	CAGAATTTCTGAAGACAAGTTTGTCTCTAAAGTTAAGATGGAGATCGTGGTTAGCAAG	470
DB	365	CTGAAATTTCTGAGGACAAATTTGTCTCTAAAGTTAATAATGGAGATGTGGTGTAGCAAG	424
QY	471	ACCAGGTTGAGGATGTTTATAGAAAAATCATTCAGGAGGCGAAGAACTGGAGAGATTGGAG	530
DB	425	ACCAAGTTGAAGCAGTAATTTGACAAGATCAATGAGTGGCAAGGACTGGAGAGATTGGTG	484
QY	531	ACGCAAGATTTTCTGCTGCTGTTTTCAGATGTAAATAGAGTCCGACATCGTGCAGCGGG	590
DB	485	ATGSCAAGATCTTTTTGGTGGCCATATCGGATGTGATAAGAGTTTCGGACCGGTGAACGTG	544
QY	591	GTGATAGGCTGAGAGGATCACGAGGGCGGATCTGACATGAGTACTTCTCTGTTGACTGC	650
DB	545	GAGAGATGGCTGAGAGATGACTGGTGGCGAACTGATATGTCCTACTACAACACCGCCTG	604
QY	651	TGTGACCAAGCAATAT	665

Db 605 CTTGAGTCCAAATAT 619

RESULT 3
 BQ798001/C
 LOCUS
 DEFINITION
 EST 6939 Ripening Grape berries Lambda Zap II Library Vitis
 vinifera cDNA clone RT093C02 3', mRNA sequence.
 ACCESSION
 BQ798001
 VERSION
 BQ798001.1 GI:22012967
 KEYWORDS
 EST
 SOURCE
 Vitis vinifera
 ORGANISM
 Vitis vinifera
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; Vitaceae; Vitis.

REFERENCE
 1 (bases 1 to 723)
 Abbal, P., Agasse, A., Ageorges, A., Atanasova, R., Barrieu, P.,
 Couture, C., Dedaldechamp, F., Delrot, S., Glissant, D., Grimplet, J.,
 Hamdi, S., Romieu, C. and Terrier, N.
 Generation of Expressed Sequence Tag from Grape Berry (skin, pulp
 or seeds) at Various Developmental Stages
 Unpublished (2002)
 Contact: Romieu C.
 Unité de Recherche des Produits de la Vigne
 Institut National de la Recherche Agronomique
 2, place Viala, 34 060 Montpellier Cedex 01, France
 Tel.: 00-33-(0)4-99-61-28-62
 Fax: 00-33-(0)4-99-61-28-57
 Email: romieu@ensam.inra.fr
 Seq primer: TV.

FEATURES
 Location/Qualifiers
 1..723
 /organism="Vitis vinifera"
 /mol_type="mRNA"
 /cultivar="Shiraz"
 /db_xref="taxon:29760"
 /clone="RT093C02"
 /dev_stage="ripening stage"
 /clone_lib="Ripening Grape berries Lambda Zap II Library"
 /notes="Organ: Fruit; Vector: Lambda Zap II; Site 1: Eco
 RI; Site 2: XhoI; Oriented library, construction described
 in Generation of ESTs from grape Berry (skin, pulp or
 seeds) at various developmental stages by Terrier, N.,
 Ageorges, A., Abbal, P., Romieu, C. in J. Plant Physiol. 158
 (12): 1575-83 2001"

ORIGIN
 Query Match 33.5%; Score 300.8; DB 13; Length 723;
 Best Local Similarity 76.0%; Pred. No. 9.9e-36;
 Matches 371; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 152 CTTAGACATCTCGGTTTCTCACTTAACACCGCGTCAAGCGCGTAAGATATGCCCCC 211
 Db 723 CTCAGAGATCTCGAATTCGAGTTTCAGTTTATTTGGCTTAACCGCTCAAGAGATGGATCT 664
 QY 212 GTCTTCTCTGTGATTAATGCGCAAGCTCGCTGACTACATTCCTGATGCTTAATTTCTAC 271
 Db 663 GTTCTTCCCATTTGACAGGCTCGAGGCTTCCAGATTATACCCAGAGCGCGCAGTTTAC 604
 QY 272 AAGTGAAGCAATTCACGCGCTCGCGCTCTCGCAAGTTTCTCGGCTTCTCAAAA 331
 Db 603 AAGTTGAAGCGATCCCTGAGCGCCCTGCGGAATCCAGCAGGTTTCTTGGCTTCTGAAA 544
 QY 332 ATTGGTATTCAGGTGTTACTGTTTCTGATCTCGAGGTTTGGTGTCTCAAGGTGGTCA 391
 Db 543 ATGGTATTCGTGTGTACTGTTTCTGATGCTCGGGGCTTGGTGTCTCAAGCGGTTCG 484
 QY 392 ACTGAGAGCGCGCGCTCAGAAATTTCTGAAGACAGATTTGTGCTAAAGTTAAGATG 451
 Db 483 CCTGAAGACAGCGCAGGCTCTGAATTTTCTGAGGACAAATTTGTTGCTAAAGTCAAAATG 424
 QY 452 GAGATCGTGGTTAGCAAGACCGAGTTGAGGATGTTATAGAAAAATCATTTGAGGAGCA 511

Db 423 GAGATTGTGTGAGCAAGACCGAGTTGAAGCAAGTAATTTGACAAGATCAATGAGGTGGCA 364

QY 512 AGAATCTGGAGAGATTGGAGACCGCAAGATTTCTTCTGCTGCTGTTTTCAGATGTAATAAGA 571
 Db 363 AGCACTGGAGAGATTGGTGTGATGCGCAAGATCTTTTGGTGGCCCATATCGGATGTGATAAGA 304

QY 572 GTCCGCACTGGTGGAGCGGGTGTATAGGCTGAGAGATGACAGAGAGCGGATCTGACATG 631
 Db 303 GTTCGACCGGTGAACGTGAGAGCTGGCTGAGAGATGACGGGTGGCGCACTGATATG 244

QY 632 AGTACTTC 639
 Db 243 TCCACTAC 236

RESULT 4
 BE239550
 LOCUS
 DEFINITION
 EST403599 MHRP- Medicago truncatula cDNA clone pMHRP-28F3, mRNA
 sequence.
 ACCESSION
 BE239550
 VERSION
 BE239550.1 GI:9055733
 KEYWORDS
 EST
 SOURCE
 Medicago truncatula (barrel medic)
 ORGANISM
 Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.

REFERENCE
 1 (bases 1 to 597)
 Harrison, M.J., Town, C.D., Bowman, C.L., Craven, M.B., Hansen, T.S.,
 Holt, I.E., Cho, J. and Fraser, C.M.
 ESTs from phosphate-starved roots of Medicago truncatula
 Unpublished (2000)
 Contact: Maria J. Harrison
 Plant Biology Division
 The Samuel Roberts Noble Foundation
 2510 Sam Noble Parkway, Ardmore, OK 73401, USA
 Tel: 580-223-5810
 Fax: 580-221-7380
 Email: mjharrison@noble.org
 The Samuel Noble Roberts Foundation: N265417e
 TIGR sequence name: MTHAC26TK
 More information is available at:
 http://chrysis.tamu.edu/medicago
 Seq primer: SKmod (CTA GAA CTA gct gat CC).

FEATURES
 Location/Qualifiers
 1..597
 /organism="Medicago truncatula"
 /mol_type="mRNA"
 /cultivar="A17"
 /db_xref="taxon:3880"
 /clone="pMHRP-28F3"
 /tissue_type="roots"
 /dev_stage="phosphate-starved"
 /lab_host="Xl0LR"
 /clone_lib="MHRP-"
 /notes="vector: Bluescript SK-; Site 1: EcoRI; Site 2:
 XhoI; At the trifoliolate stage, M. truncatula plants were
 transplanted to phosphate-free sand and grown for a
 further 30 days. During this period, they were fertilized
 twice weekly with 1/2 Hoaglands solutions containing 20uM
 potassium phosphate. cDNA was prepared from polyA+
 enriched RNA. The cDNA was directionally ligated into the
 Unizap XR vector from Stratagene and packaged using
 Gigapack III Gold packaging extracts. Plasmids containing
 cDNA inserts were excised from the recombinant lambda-Zap
 phage using Ex-assist helper phage and propagated in
 XL0LR cells."

ORIGIN
 Query Match 33.3%; Score 298.6; DB 10; Length 597;


```

362 GTTCGAGGTTTGGTGTCTCAAGGTGGTCTCACTGAGAGCGAGGCGGCTCAGAATTTTCT 42
Db      GTCAAGGGGTTTGGTGTCTCAAGGTGGTCTCAAAAGAGAGAGCGAGGAGGCTCCGAATTTTCT 415
QY      GAAGACAAGTTTGTGTCTAAAGTTAAAGATGGAGATCGTGGTTAGCAAAAGACCAGGTTGAG 481
Db      GRAGACAATTTTGTGTCCCAAGTTTAAATAGAAATAGTGGTGAGAAAAGACCAGGTTGAG 355
QY      GATGTTTAGAAAAAATCATTTGAGGAGGCAAGAACTGAGAGATTTGGAGACGCGAAGATT 541
Db      GCAGTGTGAATAACAAATTTATGGACGCGCAAGAACTGGGAGATTTGGTGATGCGCAAAATT 295
QY      TTCTTGCTGCTGCTTTTCAGATGTAATAAGAGTCCGCACTGGTGAGCGGGGTGATAAGGCT 601
Db      TTCTTGATCCCTGTATCTGTATCTGAATTAAGATTCGCAAGGTGAGCTGGGAGCGAGCT 235
QY      GAGAGGATGACAGGAGGCGGATCTGAC 628
Db      GAGAGGATGCTGGGGGACTAACTGAC 208

RESULT 7
BF645522
LOCUS   BF645522
DEFINITION   BF645522 612 bp mRNA linear EST 20-DEC-2000
            clone NF036F08EC 5', mRNA sequence.
ACCESSION   BF645522
VERSION     BF645522.1 GI:11910651
KEYWORDS    EST.
SOURCE      Medicago truncatula (barrel medic)
ORGANISM    Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1 (bases 1 to 612)
AUTHORS    Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
            Flores,H.R., Iman,J.T., Weller,J.W. and May,G.D.
TITLE      Expressed Sequence Tags from the Samuel Roberts Noble Foundation -
            Center for Medicago Genomics Research
JOURNAL     Unpublished (2000)
COMMENT     Contact: Dixon RA
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73402, USA
            Tel: 580 221 7302
            Fax: 580 221 7380
            Email: radixon@noble.org
            Insert Length: 612 Std Error: 0.00
            Plate: 036 row: F column: 08
            Seq primer: TCACACAGGAACACAGCTATGAC.
FEATURES             Location/Qualifiers
             1..612
                /organism="Medicago truncatula"
                /mol_type="mRNA"
                /db_xref="taxon:3880"
                /clone="NF036F08EC"
                /tissue_type="Cell cultures derived from root tissues"
                /dev_stage="Cell suspensions were subcultured every 14
                days. Cells were induced six days after subculture"
                /clone_lib="Elicited cell culture"
                /note="vector: Lambda Zap; Cells were induced with yeast
                cell wall extracts equivalent to 50ug/ml glucose in the
                final concentration. Samples were taken at 0.5, 1, 12 and
                24 hours after induction. Equal amounts of RNA from each
                time point were pooled and used for mRNA isolation."

```

62 ACTGCGAACTGGCTGCTCACTCCTCTCTCACTTCTAATTAACATCAAGAAGAAATTCCT 121
122 GTTTTGAATTCAGTTGTTTGGTTCAGAGCTTAGACATCTCGTCTTCTCACTTTAAC 181
104 TTTTCAAGCTTCAGCTCATTCGCAAGCTTTCGAGATTC-----TTCATCGCAAT 157
182 ACCGCGTCAAGCGCTAAGATATGCCCCCTCTCTCTGTGATTAATGCCCCAAAGTCG 241
158 GTGCTCTTAAAGCAAGCAAGTATGATCGATTCTTCCAAATCAGAGCTCAAAACCTT 217
242 CCGTACTACATTCCTGATGCTAAATCTTACAAGTGAAGCAATCTCAGGCCCTGCGGA 301
218 CCGTACTATGTTCTGAATCCAGTTTACAAAGTTGAAGCCATCTCAGGCCATGAGA 277
302 GTCTCGCAAGTTTCTCGGCTTCTTAAATTCGTAATTCGAGTGTCTGATTTCTGAT 361
278 ATCCCTCAGGTTTCTCGGCTTCTTGAATTCGTAATTCGAGTGTCTGATTTCTGAT 337
362 GTTCGAGTTTGTGCTCAAGTGTGTTCAACTGAGAGCGGCGCTCAGAAATTTCT 421
338 GTCAAGGTTTGTGCTCAGGTTGCTCAAAAGATAGGCGAGGAGCTCCGAAATTTCT 397
422 GAAGCAAGTTTGTGCTTAAAGTTAAGTGAAGTGTGTTAGCAAGACCAAGTTTCTGAG 481
398 GAAGCAATTTTGTGCTCAAGTTTAAATTCGTAATTCGAGTGTCTGATTTCTGAT 457
482 GATGTTAAGAAATCATTCAGGAGGCAAGACTGGAGAGATTCGAGAGCGCAAGATT 541
458 GCAGTGAATAAATAATTCAGGAGCAGCAAGACTGGAGAGATTCGAGAGCGCAAGATT 517
542 TTTCTGCTGCTTTCAGATGTAATAAGAGTCCGCACTGTGAGCGGGGTGATTAAGCT 601
518 TTTGATCCCTGATCTGATGATTAAGATTCGAGAGGAGTGTGATGCGCAATTTCT 577
602 GAGAGGATGACAGGAGGCGGATCTGAC 628
578 GANAGGATGGCTGGGGGACTACTGAC 604

RESULT 8
CF920467
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

CF920467 759 bp mRNA linear EST 05-NOV-2003
gmthrw3-14 G03_1 019 Soybean root hair subtracted cDNA library
gmthrw3 Glycine max cDNA, mRNA sequence.
CF920467
CF920467.1 GI:38191261
EST.
Glycine max (soybean)
Glycine max
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 759)
Scheffler,B.E., Huang,S., Liu,X., Nguyen,H., Duke,M. and Stacey,G.
Expressed sequence tags from soybean root hair subtractive cDNA
library
Unpublished (2003)
Contact: Gary Stacey
University of Missouri
108 Waters Hall, Columbia, MO 65211, USA
Tel: 573-884-4752
Fax: 573-882-0588
Email: stacey@missouri.edu
Single pass sequence
Seq primer: T7.
Location/Qualifiers
1. .759
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams 82"

RESULT 9
C81819
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

C81819 866 bp mRNA linear EST 14-AUG-2003
C81819 Citrus unshiu juice sac and pulp segment maturation stage
Citrus unshiu cDNA clone pcMFR01.17-061, mRNA sequence.
C81819
C81819.1 GI:33650789
EST.
Citrus unshiu
Citrus unshiu
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Sapindales; Rutaceae; Citrus.
1 (bases 1 to 866)
Omura,M.
Citrus unshiu juice sac and pulp segment maturation stage
Unpublished (2003)
Contact: Mitsuo Omura
Department of Citrus Research
National Institute of Fruit Tree Science, Okitsu
Okitsu 485-6, Shimizu, Shizuoka 424-0292, Japan
Tel: 81-543-69-7108
Fax: +81-543-69-2115

Query Match
Best Local Similarity 31.7%; Score 284.2; DB 14; Length 759;
Matches 367; Conservative 72.7%; Pred. No. 3.8e-35; Indels 0; Gaps 0;
/db xref="taxon:3847"
/tissue type="root hairs"
/clone_lib="Soybean root hair subtracted cDNA library
gmthrw3"
/notes="Organ: root hairs; Vector: PCR2-1 Topo; cDNA clones
generated from soybean root hair tissue treated with
Bradyrhizobium japonicum for 3 hours."

167 TTTTCTCATTTAACACCGCGCTCAAGCGCTAAGATATGCCCCCTCTCTGTGATTAATGCCCCAAAGTCG 226
127 TCTCTCAGCCCAATGCTGCTTAAGAGCGAGTAAATGGAACAATTTCTTCCCAATC 186
227 AATGCCCAAGCTCGCTGACTACATTCCTGATCTAAATCTCAAAAGTGAAGCAATTTCT 286
187 AGAGCCCAAGATCTTCCAGATTATGTTCCMAAATCCGAGTTTTCACAAAGTAGAAGCGATT 246
287 CTCAGGCCCTCGGAGCTCTCCAGATTTCCTCGGCTTTCCTAAGATTCGTAATTCGAGT 346
247 CTCAGGCCATCGGAGTTTCCCGAGTTTCTCGGCTTTCCTGAAATTCGGAATTTCTGCT 306
347 GTTACTGTTTCTGATGTTTCGAGGTTTTCGCTCAAGTGTGTTCAACTGAGAGGCGAGGC 406
307 GTCTCTGATCTGATGTTTCAGGGCTTTCGCTCAGGCTGTTTCAAAAGAGAGCGAGCA 366
407 GCTCAGAAATTTCTGAAAGCAAGTTTGTGCTAAAGTTAAGATGAGAGATTCGTTGTTAGC 466
367 GCTCCGAATTTTCAGAGACAATTTTGTGCCAAAGTTTAAATTCGAGTAGTGTGTTGAGA 426
467 AAAGCCAGGTTTGAAGATGTTTATAGAAAATTCATTGAGAGGCGAGCAACTCGAGAGATT 526
427 AAGGACCAAGTTTGAAGCAATTTTATGACAAATTTATTTAGAGAGGCAAGAACTCGGAGATT 486
527 GAGAGCGGCAAGATTTTCTGCTGCTGCTTTCAGATGTAATTAAGATTCGCACTGCTGAG 586
487 GGTGATGGCAAAATTTTCTTGATCCCATCTCAGATGTTTATGAATTCGACAGGTGAA 546
587 CGGGGTGATAAGCTGAGAGGATGACAGAGGGCGATCTGACATGATGATCTTCTGCTGCA 646
547 CGTGGGAGCAGCAGCTAGGATGACTGGGGGCGGAGTGCATGTTATCTGCTGATGA 606
647 CTGCTGTGACCAATATAGCAATT 671
607 CTGCTCGAATCAAGATTAAGCAATT 631

Email: x0145039a@frc.go.jp.
 Location/Qualifiers
 1. .866
 /organism="Citrus unshiu"
 /mol_type="mRNA"
 /cultiivar="Miyagawa-Wase"
 /db_xref="taxon:55188"
 /clone="pcwPrM01.17-061"
 /tissue type="juice sac and pulp segment"
 /dev stage="maturation stage"
 /clone lib="Citrus unshiu juice sac and pulp segment maturation stage"

FEATURES
 source
 Query Match 31.5%; Score 283; DB 13; Length 866;
 Best Local Similarity 71.8%; Pred. No. 5.3e-35;
 Matches 484; Conservative 0; Mismatches 143; Indels 47; Gaps 7;

ORIGIN
 QY 8 CGGCGCTCTAGAACTAGTGGATCCCGGGCTGCAGGAATTCGGCAGCAG-----58
 DB 20 CGNCCGCTCTAGAACTAGTGGATCCCGGGCTGCAGGAATTCGGCAGCAGATCGTTAA 79
 QY 59 -----GCTACTGCCAAACTGGGCTTGCTCACTCTCTTCATTCTA 98
 DB 80 CTGCGAGACCGTAATGGCTGCAATGGCAACCCAGCTCGTTACTCGCTCAACTTCT 139
 QY 99 ATAAATCAAGAAAGAAATTCCTGTTTGGATTTCAGTTTGGTTTGTCCAGAGCTTAGAC 158
 DB 140 CTTCTCATCTCAAGAAATTCCTCTTGTAAATTCGGCTCCGTCGGCGCAAGTTTGAGAA 199
 QY 159 ATTCTCGGTTTCTCACTTTACACCGCGCTCAAGCGGTAAGATATGCCCGCTCGTTC 218
 DB 200 TTTCGAGCTTTCTAG-----GCTCAATGCAAGAAATGATCAAACTTTC 247
 QY 219 CTGTGATTAATGCCAAAGCTCGCTGACTACATTCCTGATGCTAAATCTACAAAGTGG 278
 DB 248 ATGTTATCAGAGCCAGAGCTCACTGATTAATTCCTGATTTCTAAGTTTACAAAGTGG 307
 QY 279 AAGCAATTCAGGCGCTCGGAGTCTCGAAGTTTCTCGGCTTCTGCTAAATTTGGTA 338
 DB 308 AAGCCATTCAGGCGCTCGGAGTCCAGAGTTTCTTCGGCTTTGCTGAATATGGGTA 367
 QY 339 TTGAGGCTTACTGTTCTGATGTCGA--GGTTTGGTCTCAAGTGCTTCAAC--TGA 396
 DB 368 TCGTGGTGTACGGTTCTGATGTTGATGTTTGGTCTCAAGTGCTCAAGTGCTCAACAGGA 427
 QY 397 GAGGAGGCGGCT--CAGAAATTTCTGA--AGACAAGTTTGTGCTAAAGTTAAGATGA 453
 DB 428 AAGGCATGTGCTCCCTGATTTCTGAGGGAATAATTTGTCNAAAGTGAAGATGA 487
 QY 454 GATCGTGTAGCAAGACCAAGTGTAGGATGTTATAGAAAATCATTTGAGGAGGCAAG 513
 DB 488 GATTGTGG--TAACAAGACCAAGTGAAGAGTAATTGATAAAATTTATGAGGAGGCGCAG 546
 QY 514 ACTCGAGAGATGGAGACGCGCAAGATTTCTTGTGCTGCTGTTTTCAGATGTAATAGAGT 573
 DB 547 GACTCGGAATTTGGTACGCGCAAGATTTTGTGTTACCTGTCTCAGATGTCATAGAGT 606
 QY 574 CCGCACTGTGTGAGCGGCTGATAGGCTGAGAGGATGACAGGAGGCGATCTGCATGAG 633
 DB 607 TCGNACTGTGACAGAGTGAAGAGCTGAAGAGTGAAGGATGCGAGGAGTGTGCGACATATC 666
 QY 634 TACTTCTGTTGAC 647
 DB 667 TTCAGTGTGTGAC 680

RESULT 10
 BE823431/c
 LOCUS
 DEFINITION GM700019823E12 Gm-r1070 Glycine max cDNA clone Gm-r1070-7536 3', mRNA sequence.
 BE823431
 ACCESSION

BE823431.1 GI:10255665

EST.

Glycine max (soybean)

Glycine max

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.

1 (bases 1 to 677)

Vodkin,L., Keim,P., Shoemaker,R., Retzel,E., Khanna,A., Corryell,V.,
 Erpellding,U., Raph,C., Shoop,E., Pardini,J., Liu,L. and Lewin,H.
 A Functional Genomics Program for Soybean (NSF 9872565)
 Unpublished (1999)

JOURNAL

Other ESTs: AW471508 corresponding to Gm-cl029-383 (5')

Contact: Vodkin, L.O., PI, A Functional Genomics Program for
 Soybean (NSF 9872565)
 Lewin, H. A., Director, Keck Center for Comparative and Functional
 Genomics

University of Illinois

Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA

Tel: (217) 244-6147

Fax: (217) 333-4582

Email: l-vodkin@uiuc.edu

This clone is available through: Genome Systems, Inc. 4633 World
 Parkway Circle St. Louis, Missouri 63134. For further information
 call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)
 427-3324 or contact: clones@genomesystems.com or info@genome
 systems.com web site: www.genomesystems.com

Seq primer: 5'-TTTTTTTTTTTTTTT(A/C/G)-3'.

Location/Qualifiers

FEATURES

source

1. .677

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clone lib="Gm-r1070"

/clone lib="Gm-r1070"

The library Gm-r1070 is a sequence-driven, reracked
 set of 9,216 clones selected from cDNA libraries from
 various tissues and stages of development of soybean that
 represent 2,639 sequences from immature cotyledons, 1,770
 from immature seed coats, 3,938 from flowers, and 869
 from young pods. The 5' ESTs of the source clones from
 the different libraries was used to select singletons, or
 a representative of each contig, which were reracked to
 form library Gm-r1070. The cDNA clones of the reracked
 Gm-r1070 library were then sequenced at the 3' end. The
 contig analysis to select unique genes was performed by
 the laboratory of Ernest Retzel, Center for Computational
 Genomics and Bioinformatics, University of Minnesota,
 http://www.cbc.umn.edu/ResearchProjects/Soybean/index.html
 . Reracking was performed by Genome Systems, St. Louis,
 http://www.genomesystems.com, and 3' sequencing by the
 Keck Center for Comparative and Functional Genomics,
 University of Illinois.

Note: The
 corresponding 5' EST from each clone in the Gm-r1070
 library is listed in the 'OTHER EST' field. The detailed
 information on the source library for each clone can also
 be obtained by referring to the Genome Systems clone ID of
 the original cDNA library that is also listed under
 'OTHER EST'."

ORIGIN

Query Match 31.0%; Score 277.8; DB 10; Length 677;

Best Local Similarity 72.2%; Pred. No. 4e-34;

Matches 345; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

QY 194 CGCGTAAGATATGCCCGCTCGTTCTGTGATTAATGCCCAAGCTCGCTGACTACATT 253

DB 675 CACAGATTAATNN 616

QY 254 CCTGATGCTAAATTTCTACAAAGTGAAGCAATTTCTAGCGCTCGCGAGTTCGCAAGTT 313

615	Db	CCAAAATCCGAGTTTTCAAAAGTAGAAGCGATTCTCAGNCAATGCGGAGTTCNNNNGTT	556
314	Qy	TCCTCGGCTTTCGCTAAAAATGGTATTCCAGAGTGTTACTGTTTCTTGATGTTCCGAGTTTT	373
555	Db	TCTCGGCTTGTTCAAAATGGGAATTCGTGGTGTCACTGTATCTGATGTGAGGGCTTT	496
374	Qy	GGTCTCAAGGTGGTTTCAACTGAGAGGCAAGGGCGGCTCAGAAATTTCTGAACACAAGTTT	433
495	Db	GGTCTCAGGCTGGTTCAAAAAGAGAGCGAGGAGGCTCCGAAATTTCTGAAGACACAATTT	436
434	Qy	GTTCTAAAGTTTACATCGAGATTCGTGTTTACCAAGACCAGCTTGAGGATGTTATAGAA	493
435	Db	GTTGCCAAAGTTTAAATCGAAGTAGTGGTGGAAGAGGCCAGCTTGAGCGAGTTATGAC	376
494	Qy	AAAATCATTTGGAGGCGCAAGAACTCGAGAGATTGGAGACGGCAAGATTTTCTGCTGCT	553
375	Db	AAAATTTTGAAGGCGCAAGAACTCGGGGAGATTGGTGAATGGCAAAATTTTCTTGATCCCC	316
554	Qy	GTTTCAGATGTAATAAGAGTCCGCACTGCTGAGCCGGGGTGATAAGCTCGAGAGGATGCA	613
315	Db	ATCTCAGATGTTATAAGAAATTCGACAGGTGAACGTGGGAGCAGGCGAGCTAGGATGACT	256
614	Qy	GGAGGGCGATCTGACATGAGTACTTCTGCTTCACTGCTGTGACCCAGCAATATAGCAAT	671
255	Db	GGGGGCCCAAGTGACATGTTATCTGCTGTATGACTGGTTGGAAATGAAGATGAAGATT	198

RESULT 11	BM113473	799 bp	linear	EST 10-MAR-2003
LOCUS	EST563009	potato roots	So1anum tuberosum	cDNA clone cFRO18M20.5
DEFINITION	end, mRNA sequence.			

KEYWORD	SOURCE	ORGANISM
ESI.		<i>Solanum tuberosum</i> (potato)
		<i>Solanum tuberosum</i>
		<i>Solanum tuberosum</i>
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterid; lamiales; Solanales; Solanaceae; Solanum.

FEATURES
source

```

Location/Qualifiers
1. 799
  /organism="Solanum tuberosum"
  /mol_type="mRNA"
  /cultivar="Kennebec"
  /db_xref="axon:4113"
  /clone="CPRO18M20"
  /tissue_type="roots"
  /dev_stage="in vitro grown stem cuttings"
  /lab_host="SOLR"
  /clone_lib="potato roots"
  /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Cornell University, Tanksley lab;
sequencing: The Institute for Genomic Research. Roots were
isolated from in vitro grown stem cuttings on CM medium.
Roots were isolated two weeks after placing the stem
cuttings from in vitro grown plants on medium."

```

ORIGIN

FEATURES	SOURCE
-----------------	---------------

```
seq primer: 13:
Location/Qualifiers
1..711
/organism="Lycopersicon esculentum"
/mol type="mRNA"
```

Query Match	30.8%	Score 276;	DB 12;	Length 799;
Best Local Similarity	70.4%	Pred. No. 6.9e-34;		
Matches 369: Conservative	0;	Mismatches 155;	Indels 0;	Gaps 0;

RESULT 12
BI932336

LOCUS	B1932336	711 bp	linear	EST 18-Oct-2000
DEFINITION	BT552225 tomato flower, 8 mm to preanthesis buds Lycopersicon esculentum cDNA clone CTC22H10 5' end, mRNA sequence.			
ACCESSION	B1932336			
VERSION	B1932336.1	GI:16246808		
KEYWORDS	EST.			
SOURCE	Lycopersicon esculentum (tomato)			
ORGANISM	Lycopersicon esculentum			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon. 1 (bases 1 to 711)			

TITLE	JOURNAL	COMMENT
1. The Role of the Teacher in the Classroom	Journal of Educational Research	1965, Vol. 68, No. 1, pp. 1-10
2. The Impact of Technology on Education	Journal of Educational Technology	1978, Vol. 1, No. 2, pp. 1-10
3. The Importance of Parental Involvement	Journal of Educational Research	1985, Vol. 88, No. 3, pp. 1-10
4. The Role of the School in the Community	Journal of Educational Research	1992, Vol. 95, No. 4, pp. 1-10
5. The Impact of Teacher Education on Student Achievement	Journal of Educational Research	1998, Vol. 101, No. 5, pp. 1-10
6. The Role of the Teacher in the Classroom	Journal of Educational Research	2005, Vol. 108, No. 6, pp. 1-10
7. The Impact of Technology on Education	Journal of Educational Technology	2012, Vol. 3, No. 1, pp. 1-10
8. The Importance of Parental Involvement	Journal of Educational Research	2018, Vol. 121, No. 2, pp. 1-10
9. The Role of the School in the Community	Journal of Educational Research	2020, Vol. 123, No. 3, pp. 1-10
10. The Impact of Teacher Education on Student Achievement	Journal of Educational Research	2022, Vol. 125, No. 4, pp. 1-10

TITLE Generation of ESTs from tomato flower tissue, buds 8 mm -
preanthesis
JOURNAL Unpublished (2001)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
This clone is available through the Clemson University Genomics
Institute

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Computer Ltd.

OK nucleic - nucleic search, using sw model

Run on: May 27, 2004, 13:35:34 ; Search time 1738.01 Seconds
(without alignments)
14663.718 Million cell updates/sec

Title: US-09-756-541-15

Perfect score: 588

Sequence: 1 ATGGCGGGCTCAATGACGAA.....GTGATATGCTTTACCGTCT 588

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 6

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: gb.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.fam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	588	100.0	588	6	AR125591	AR125591 Sequence
2	588	100.0	591	6	AX507438	AX507438 Sequence
3	588	100.0	591	8	BT005209	BT005209 Arabidops
4	588	100.0	817	6	AR125589	AR125589 Sequence
5	588	100.0	844	8	AF095455	AF095455 Arabidops
6	285	48.5	534	6	AR125592	AR125592 Sequence
7	285	48.5	840	8	AF095454	AF095454 Ricinus c
8	285	48.5	897	6	AR125590	AR125590 Sequence
9	250.8	42.7	796	8	AY027892	AY027892 Medicago
10	245	41.7	947	8	AY442185	AY442185 Lycopersi
11	209.6	35.6	1085	8	PPI489604	AJ489604 pinus pin
12	199.8	34.0	120185	8	AC007138	AC007138 Arabidops
13	199.8	34.0	198220	8	ATCHRIVS	AL161493 Arabidops
14	199.4	33.9	902	8	AK068407	AK068407 Oryza sat
15	199.4	33.9	902	8	AK099152	AK099152 Oryza sat
16	109.8	18.7	384	1	FDGLNBPT	AE017165 Prochloro
17	107.4	18.3	300143	1	AE017165	AE017165 Prochloro
18	106.8	18.2	164921	8	AF022186	AF022186 Cyanidium
19	106	18.0	2493	1	ASP251822	AT251822 Anabaena
20	106	18.0	341880	1	AP003589	AP003589 Nostoc sp
21	102	17.3	262202	1	BX572094	BX572094 Prochloro
22	101.6	17.3	191028	8	PPU38804	U38804 Porphyra pu
23	98.6	16.8	4348	1	AF079137	AF079137 Synchoco
24	98	16.7	685	1	AF017419	AF017419 Nostoc pu
25	97	16.5	383	1	SYOGLNB	M62447 Synchococcc
26	95.8	16.3	339	1	PMA271089	AJ271089 Prochloro
27	95.8	16.3	11771	1	AE000674	AE000674 Aquifex a
28	94.2	16.0	349746	1	BX572099	BX572099 Prochloro
29	94	16.0	299350	1	AP005370	AP005370 Thermosyn
30	93.6	15.9	302050	1	AP005568	AP005568 Gloeobact
31	93.6	15.9	302320	1	AP005081	AP005081 Vibrio pa
32	93.2	15.9	300045	1	AE016803	AE016803 Vibrio vu
33	92.8	15.8	96109	6	AR408756	AR408756 Sequence
34	92.8	15.8	96109	6	AX067460	AX067460 Sequence
35	91.6	15.6	247950	1	AP005340	AP005340 Vibrio vu
36	89.8	15.3	11371	1	AE006236	AE006236 Pasteurel
37	89.4	15.2	351	6	AR378407	AR378407 Sequence
38	89.4	15.2	300732	1	AE016802	AE016802 Vibrio vu
39	89	15.1	1622	1	AF120107	AF120107 Synchoco
40	88.4	15.0	452	1	SSGLNBP	X97496 Synchococyst
41	88.4	15.0	130001	1	D90915	D90915 Synchococyst
42	86.8	14.8	298950	1	AP004597	AP004597 Oceanobac
43	86.2	14.7	248650	1	AP005341	AP005341 Vibrio vu
44	85.8	14.6	375	6	AR318761	AR318761 Sequence
45	85.4	14.5	405	6	AR376903	AR376903 Sequence

ALIGNMENTS

RESULT 1	AR125591	AR125591	588 bp	DNA	linear	PAT 16-MAY-2001
LOCUS	Sequence 15 from patent US 6177275.					
DEFINITION	Sequence 15 from patent US 6177275.					
ACCESSION	AR125591					
VERSION	AR125591.1	GI:14111653				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 588)					
AUTHORS	Coruzzi, G.M., Lam, H.-M. and Hsieh, M.-H.					
TITLE	Plant nitrogen regulatory P-III genes					
JOURNAL	Patent: US 6177275-A 15 23-JAN-2001;					
FEATURES	Location/Qualifiers					

```

source
ORIGIN
Query Match 100.0%; Score 588; DB 6; Length 588;
Best Local Similarity 100.0%; Pred. No. 2.8e-137;
Matches 588; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGGCGGCGTCAATGACGAAACCCATCTCAATTAACCTTCTCTCGGTTTCTATTCTGATCGA 60
Db
1 ATGGCGGCGTCAATGACGAAACCCATCTCAATTAACCTTCTCTCGGTTTCTATTCTGATCGA 60
QY
61 AAGAACATTGCTTCTCTGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 120
Db
61 AAGAACATTGCTTCTCTGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 120
QY
121 TCTTGCCCTCGATTTGGTGCACAAAGTCAACGAGTAATAACAGTCTGCTGTTTACCTGCTGTT 180
Db
121 TCTTGCCCTCGATTTGGTGCACAAAGTCAACGAGTAATAACAGTCTGCTGTTTACCTGCTGTT 180
QY
181 AGTGCCCAAAATATCTTCTGATTATATTCAGACTCGAAATTTTACAGGTTGGAAGCAATT 240
Db
181 AGTGCCCAAAATATCTTCTGATTATATTCAGACTCGAAATTTTACAGGTTGGAAGCAATT 240
QY
241 GTCAGACCATGAGAAATCCAGCAAGTTTCATCGCTTTCATCGAAATTCGAGATTCGAGGT 300
Db
241 GTCAGACCATGAGAAATCCAGCAAGTTTCATCGCTTTCATCGAAATTCGAGATTCGAGGT 300
QY
301 GTTACTGTTTCTGATGTCAGAGGTTTGGTGCACAAAGGAGTCTTACCGAGAGACACGGT 360
Db
301 GTTACTGTTTCTGATGTCAGAGGTTTGGTGCACAAAGGAGTCTTACCGAGAGACACGGT 360
QY
361 GGCCTCTGAGTTCTCGGAGACAAATTTGTTGCTAAAGTTAAGTGAAGATCGTTGTTAAG 420
Db
361 GGCCTCTGAGTTCTCGGAGACAAATTTGTTGCTAAAGTTAAGTGAAGATCGTTGTTAAG 420
QY
421 AAAGACCAAGTGGAAATCTGTAATCAACAAATTTGTTGCTAAAGTTAAGTGAAGATCGTTGTTAAG 480
Db
421 AAAGACCAAGTGGAAATCTGTAATCAACAAATTTGTTGCTAAAGTTAAGTGAAGATCGTTGTTAAG 480
QY
481 GGTGATGGCAAGATTTTGTGCTGTCAGATGTCATGATGCTTACCGTCT 540
Db
481 GGTGATGGCAAGATTTTGTGCTGTCAGATGTCATGATGCTTACCGTCT 540
QY
541 CGTGGGGAAGACGAGAGAGATGATGCTGATATGCTTACCGTCT 588
Db
541 CGTGGGGAAGACGAGAGAGATGATGCTGATATGCTTACCGTCT 588

RESULT 3
BT005209 591 bp mRNA linear P.N 14-MAR-2003
Arabisopsis thaliana At4g01900 mRNA, complete cds.
DEFINITION
ACCESSION BT005209
VERSION BT005209.1 GI:28950698
KEYWORDS FLI CDNA.
SOURCE
ORGANISM
Arabisopsis thaliana (thale cress)
Arabisopsis thaliana
Arabisopsis thaliana
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 591)
Kim, C.J., Chen, H., Dale, J.M., Hayashizaki, Y., Hsuan, V.W.,
Chen, M.M., Chang, C.H., Karlin-Neumann, G., Kawai, J.,
Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,
Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,
Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K.,
Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and
Ecker, J.R.
Arabisopsis ORF clones
Unpublished
2 (bases 1 to 591)
Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Carninci, P.,
AUTHORS

TITLE
JOURNAL
REFERENCE
Arabisopsis thaliana
/organism="Arabisopsis thaliana"
/feature="unassigned dna"

```

Chan, M.M., Chang, C.H., Dale, J.M., Hayashizaki, Y., Hsuan, V.W., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Davis, R.W., Theologis, A. and Ecker, J.R.

Direct Submission
Submitted (14-MAR-2003) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGE (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAPL cDNAs: Kim, C.J., Chen, H., Cheuk, R., Shim, P., Bowser, L., Chan, M.M., Chang, C.M., Dale, J.M., Hsuan, V.W., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A., and Ecker, J.R.

Kim, C.J. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES

Location/Qualifiers

1..591
/organism="Arabidopsis thaliana"
/mol_type="cDNA"
/db_xref="taxon:3702"
/chromosome="4"
/clone="U23463"
/ecotype="Columbia"
/note="This clone is in pUNI 51"

CDS

1..591
/note="p II nitrogen sensing protein GLB 1"
/codon_start=1
/product="At4g01900"
/protein_id="AA063273.1"
/db_xref="GI:28950699"
/translation="MAASMTKPISTLSGYSDRKNIAFSDCISGFRHSRPSCLD
LVTKSPNSRVLVPVSAQISDIIPDSFYKVEALVPRWIOQVSSALKIGIRVY
VSDVRGFGAGSGSTERHSGSEPSDEDFVAKVMEIVVKQDVESVNTIIEGARTGEI
GDGKIFVLPSVDIVRVTGGRGEKAEKMTGDLSPS"

ORIGIN

Query Match 100.0%; Score 588; DB 8; Length 591;
Best Local Similarity 100.0%; Pred. No. 2.8e-137;
Matches 588; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGGCGTCAATGACGAACCCATCTCAATTAACCTCTCGGTTTCTATTCTGATCGA 60
DB 1 ATGCGGCGGTCATGACGAACCCATCTCAATTAACCTCTCGGTTTCTATTCTGATCGA 60

QY 61 AAGAACATTCCTCTGATGTCATTTGATTTGTTCTGATTCAGACATTCGCCACCA 120
DB 61 AAGAACATTCCTCTGATGTCATTTGATTTGTTCTGATTCAGACATTCGCCACCA 120

QY 121 TCTTGCTCGATTGGTTCACAAAGTACCGAGTAATAACAGTCTGTGTTTAACTGTGCT 180
DB 121 TCTTGCTCGATTGGTTCACAAAGTACCGAGTAATAACAGTCTGTGTTTAACTGTGCT 180

QY 181 AGTGCCCAAAATATCTTCTGATTAATTTCCAGATCCGAAATTTTCAAGGTGGAAGCAATT 240
DB 181 AGTGCCCAAAATATCTTCTGATTAATTTCCAGATCCGAAATTTTCAAGGTGGAAGCAATT 240

QY 241 GTCCAGACCATGGAGAAATCCAGAGTTTCTCGGCTTTTACTGAAAATCGGGATTCGAGGT 300

DB 241 GTCCAGACCATGGAGAAATCCAGAGTTTCTCGGCTTTTACTGAAAATCGGGATTCGAGGT 300

QY 301 GTTACTGTTTCTGATGTCAGAGGTTTGGTGCCACAGAGGTTCTACCGAGACACCGT 360
DB 301 GTTACTGTTTCTGATGTCAGAGGTTTGGTGCCACAGAGGTTCTACCGAGACACCGT 360

QY 361 GGTCTGAGTTCTCGGAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTTAAG 420
DB 361 GGTCTGAGTTCTCGGAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTTAAG 420

QY 421 AAGACCAAGTGAATCTGTAATCAACAAATTAATTAAGAGCAAGACAGGAGATTT 480
DB 421 AAGACCAAGTGAATCTGTAATCAACAAATTAATTAAGAGCAAGACAGGAGATTT 480

QY 481 GGTGATGGCAGATTTTGTGTTTGGCTGTCTCAGATGTCATTAAGATTAGACAGGTGAG 540
DB 481 GGTGATGGCAGATTTTGTGTTTGGCTGTCTCAGATGTCATTAAGATTAGACAGGTGAG 540

QY 541 CGTGGGAGAGAAAGCAGAGAGATGACTGGTGATATGCTTTCACCGTCT 588
DB 541 CGTGGGAGAGAAAGCAGAGAGATGACTGGTGATATGCTTTCACCGTCT 588

RESULT 4
ARI25589
LOCUS ARI25589 817 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 13 from patent US 6177275.
ACCESSION ARI25589
VERSION ARI25589.1 GI:14111651
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 817)
AUTHORS Coruzzi, G.M., Lam, H.-M. and Hsieh, M.-H.
TITLE Plant nitrogen regulatory P-PII genes
JOURNAL Patent: US 6177275-A 13 23-JAN-2001;
FEATURES
source 1..817
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 588; DB 6; Length 817;
Best Local Similarity 100.0%; Pred. No. 2.7e-137;
Matches 588; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGGCGTCAATGACGAACCCATCTCAATTAACCTCTCGGTTTCTATTCTGATCGA 60
DB 33 ATGCGGCGGTCATGACGAACCCATCTCAATTAACCTCTCGGTTTCTATTCTGATCGA 92

QY 61 AAGAACATTCCTCTGATGTCATTTGATTTGTTCTGATTCAGACATTCGCCACCA 120
DB 93 AAGAACATTCCTCTGATGTCATTTGATTTGTTCTGATTCAGACATTCGCCACCA 152

QY 121 TCTTGCTCGATTGGTTCACAAAGTACCGAGTAATAACAGTCTGTGTTTAACTGTGCT 180
DB 153 TCTTGCTCGATTGGTTCACAAAGTACCGAGTAATAACAGTCTGTGTTTAACTGTGCT 212

QY 181 AGTGCCCAAAATATCTTCTGATTAATTTCCAGATCCGAAATTTTCAAGGTGGAAGCAATT 240
DB 213 AGTGCCCAAAATATCTTCTGATTAATTTCCAGATCCGAAATTTTCAAGGTGGAAGCAATT 272

QY 241 GTCCAGACCATGGAGAAATCCAGAGTTTCTCGGCTTTTACTGAAAATCGGGATTCGAGGT 300
DB 273 GTCCAGACCATGGAGAAATCCAGAGTTTCTCGGCTTTTACTGAAAATCGGGATTCGAGGT 332

QY 301 GTTACTGTTTCTGATGTCAGAGGTTTGGTGCCACAGAGGTTCTACCGAGACACCGT 360
DB 333 GTTACTGTTTCTGATGTCAGAGGTTTGGTGCCACAGAGGTTCTACCGAGACACCGT 392

QY 361 GGTCTGAGTTCTCGGAGACAAATTTGTTGCTAAAGTTAAGATGGAATCGTTGTTAAG 420


```

Db      244  CCTCGCGAGTCTCGCAGTTCTCTCGGCTTGCTTAAATAATGGTATTTCGAGGTGTACT 303
Qy      307  GTTCTGATGTGAGAGGTTTGGTGCAAGAAGGTTTACCGAGAGACACGGTGGCTCT 366
Db      304  GTTCTGATGTGAGAGGTTTGGTGCTCAAGGTGGTTCAACTGAGAGGCGAGCGGCTCA 363
Qy      367  GAGTTCTCGAGACAAATTTGCTGCTAAAGTTAAGATGGAATCGTTGTAAGAAGAC 426
Db      364  GAAATTTCTGAAGAAGAAGTTTGTCTAAAGTTAAGATGGAATCGTTGTAAGAAGAC 423
Qy      427  CAAAGTGAATCTGTAATCAACAATAATTAAGGAGCAAGAGACAGGAGATGCTGAT 486
Db      424  CAGGTTGAGGATGTTATAGAAAATCAATTGAGGAGGCAAGAACTGAGAGATGGAGAC 483
Qy      487  GCGAGATTTTGTGCTGCTGTGTCAGATGTCATAGAGTTAGGACAGGTGAGGTGG 546
Db      484  GCGAGATTTTGTGCTGCTGTGTCAGATGTCATAGAGTTAGGACAGGTGAGGTGG 543
Qy      547  GAGAAGCAGAGAAGATGACTGTG 571
Db      544  GATAAGGCTGAGAGGATGACAGAG 568

RESULT 7
AF095454      AF095454      840 bp      mRNA      linear      PLN 19-NOV-1998
LOCUS      Ricinus communis PII protein mRNA, partial cds.
DEFINITION      AF095454
ACCESSION      AF095454
VERSION      AF095454.1 GI:3885940
KEYWORDS
SOURCE      Ricinus communis (castor bean)
ORGANISM      Ricinus communis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Malpighiales; Euphorbiaceae; Acalyphoideae;
Acalyphaceae; Ricinus.
1 (bases 1 to 840).
Hsieh,M.H., Lam,H.M., van de Loo,P.J. and Coruzzi,G.
A PII-like protein in Arabidopsis: putative role in nitrogen
sensing
Proc. Natl. Acad. Sci. U.S.A. 95 (23), 13965-13970 (1998)
MEDLINE      99030678
JOURNAL      9811909
PUBMED
REFERENCE      2 (bases 1 to 840)
AUTHORS      Hsieh,M.H., Lam,H.M., van de Loo,P.J. and Coruzzi,G.M.
TITLE      Direct Submission
JOURNAL      Submitted (29-SEP-1998) Biology, New York University, 100
Washington Square East, New York, NY 10003, USA
FEATURES
source
1..840
/organism="Ricinus communis"
/mol_type="mRNA"
/db_xref="taxon:3988"
<1..595
/function="putative role in nitrogen sensing"
/note="GLB1"
/codon_start=2
/product="PII protein"
/protein_id="AAC78332.1"
/db_xref="GI:3885941"
/translators="HEATKIGLLTPLHNNIKKEPFFDFSLPCPLRHSRHSFHT
AVKVPYAPVVPVINAQSPDYTPDAKFYKVRILRPMRVSOVSSALLKIGIRGVTS
DVRFGAGGSTRQGSFSEDKFVAKYKMEIVSKDQVEDVIEKLIBEARTGEIGD
GXIFLLPVSDVIRVIRGERGDKAERTMGKSDMSTA"

ORIGIN
Query Match      48.5%; Score 285; DB 8; Length 840;
Best Local Similarity 70.3%; Pred. No. 4.7e-61;
Matches 397; Conservative 0; Mismatches 165; Indels 3; Gaps 1;

Qy      7  GCGTCAATGAGAAACCCATCTCAATAACTTCTCTCGGTTTCTATTCTGATCGAAAGAC 66
Db      5  GAGGCTACTCGGAACACTGGGCTTCTCACTCTCTTCAATCTAATAACATCAAGAAGAA 64

Query Match      48.5%; Score 285; DB 8; Length 840;
Best Local Similarity 70.3%; Pred. No. 4.7e-61;
Matches 397; Conservative 0; Mismatches 165; Indels 3; Gaps 1;

```

```

Qy      67  ATTGCTTCTCTGATTCGATTTCTGATTTCTGATTCAGACATTCGCCACCATCTGC 126
Db      65  TTCCCTGTTTTGATTTCAAGTTTGTGTTTGTCCAGAGCTTAGACATTTCTCGTTCTCAC 124
Qy      127  CTCGATTTGGTTCACAAAGTCACCGAGTAATAACAGTCGTGTTTTTACCTGCTGTAGTGC 186
Db      125  TTAAACCCGCGGTCAAGCGGTAAAGATATGCC---CCGTCGTTCTGTGATTAATGCC 181
Qy      187  CAAATATCTTCTGATTAATATCCAGACTCGAAATTTTACAGGTCGGAAGCAATTTGTCAGA 246
Db      182  CAAAGCTCGCCTGACTACATTCCTGATGCTAAATTTCTCAAAAGTGAAGCAATTTCTCAGG 241
Qy      247  CCAATGAGAAATCCAGCAAGTTTTCATCGGCTTTACTGAAATCCGGATTCGAGGTGTACT 306
Db      242  CCTCGGAGTCTCGCAAGTTTCTCGGCTTTGCTTAAATTTGGTATTCGAGGTGTACT 301
Qy      307  GTTCTCTGATGTGAGAGGTTTGGTGCAAGAAGGTTTCTACCCAGAGACACCGTGGCTCT 366
Db      302  GTTCTCTGATGTTCGAGGTTTGGTGCTCAAGGTGGTTCAACTGAGAGCGAGGCGGCTCA 361
Qy      367  GAGTCTCGGAGACAAATTTGTTGCTAAAGTTAAAGTGAATCGGAATCGTTGTAAGAAGAC 426
Db      362  GAAATTTCTGAAGACAAGTTTGTGCTAAAGTTAAAGTGAATCGGATCGGTGTTAGCAAAGAC 421
Qy      427  CAAATGCAATCTGTAATCAACACAATAATTGAAGGAGCAAGCAAGAGATTTGGTGAT 486
Db      422  CAGGTTGAGGATGTTATAGAAAATCAATTCAGAGGCGCAAGAACTCGAGAGATTTGGAGAC 481
Qy      487  GCGAGATTTTGTGCTGCTGTGTCAGATGTCATAGAGTTAGGACAGGTGAGCGTGGG 546
Db      482  GCGAAGATTTCTGCTGCTGTTTCAGATGTAATAAGAGTCCGCACTGGTGGAGCGGGT 541
Qy      547  GAGAAGCAGAGAGAGTACTGTG 571
Db      542  GATAAGGCTGAGAGGATGACAGAG 566

RESULT 8
AR125590      AR125590      897 bp      DNA      linear      PAT 16-MAY-2001
LOCUS      Sequence 14 from patent US 6177275.
DEFINITION      AR125590
ACCESSION      AR125590
VERSION      AR125590.1 GI:14111652
KEYWORDS
SOURCE      Unknown.
ORGANISM      Unclassified.
REFERENCE      1 (bases 1 to 897)
AUTHORS      Coruzzi,G.M., Lam,H.-M. and Hsieh,M.-H.
TITLE      Plant nitrogen regulatory P-PII genes
JOURNAL      Patent: US 6177275-A 14 23-JAN-2001.
FEATURES
Location/Qualifiers
source
1..897
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match      48.5%; Score 285; DB 6; Length 897;
Best Local Similarity 70.3%; Pred. No. 4.7e-61;
Matches 397; Conservative 0; Mismatches 165; Indels 3; Gaps 1;

Qy      7  GCGTCAATGAGAAACCCATCTCAATAACTTCTCTCGGTTTCTATTCTGATCGAAAGAC 66
Db      5  GAGGCTACTCGGAACACTGGGCTTCTCACTCTCTTCAATCTAATAACATCAAGAAGAA 115
Qy      67  ATTGCTTCTCTGATTCGATTTCTGATTTCTGATTCAGACATTCGCCACCATCTGC 126
Db      116  TTCCCTGTTTTGATTTCAAGTTTGTGTTTGTCCAGAGCTTAGACATTTCTCGTTCTCAC 175
Qy      127  CTCGATTTGGTTCACAAAGTCACCGAGTAATAACAGTCGTGTTTTTACCTGCTGTAGTGC 186
Db      176  TTAAACCCGCGGTCAAGCGGTAAAGATATGCC---CCCGTCGTTCTGTGATTAATGCC 232

```


Query Match 41.7%; Score 245; DB 8; Length 947;
 Best Local Similarity 72.5%; Pred. No. 5.4e-51;
 Matches 317; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 135 GGTCAAAAGTCAACGAGTAATAACAGTCGTGTTTACCTGTCGTGTAGTCCCAATATC 194
 DB 178 GCTCACCTTCAACGCTGTCAAAATGCTCTCTTCCGATTATCAGAGCCCAAAATC 237

QY 195 TTCTGATTACATTCAGACTCGAAATTTTACAGGTGGAGCAATGTCCAGCAATGGAG 254
 DB 238 TCCAGATTTCGTCCTGATGCGCAAGTTTACAAAGTTGAAGCAATTTAAGACCTTGGAG 297

QY 255 AATCCAGCAAGTTTCATCGCTTTACTGAAATCGGATTCAGGTGTTTACTGTTTCTGA 314
 DB 298 AATTCACAGAGTTTCTCGGCACTACTGAAATGGGCATTCGTGGTGCACHTTTCGGA 357

QY 315 TGTGAGAGGTTTGTGTCACAAAGGAGTTCCTACCGAGAGACAGCGTGGCTCTGAGTTCTC 374
 DB 358 TGTCTGTGTTTGGCGCCCAAGTGGCTTGACTGAGAGGCAAGCTGGCTCTGAATTC 417

QY 375 GGAAGCAAAATTTGTTGCTAAAGTTAAGATGGAATCGTGTGAAGAAAGCAAGTGA 434
 DB 418 TCAAGACAGTTTGTGTCAAAAGTTAAATGGAATTTGTTGTCAGCAAAAGACCAAGTTGA 477

QY 435 ATCTGTAATCAACACATAATTTGAAGGAGCAAGGACAGAGATTTGTTGATGCGCAAGAT 494
 DB 478 AGGAGTCATTCGCAAGATTAATGAAGAGGCAAGCACTGGTGAATAGTGTGATGGAAGAT 537

QY 495 TTTTGTGTTGCTGTTGTCAGATGTCATAAGAGTTAGACAGGTGAGCGTGGGAGAAAGC 554
 DB 538 ATTCCTGACTCCCATCTCCGATTTATAGAGTTGCACTGGTGAACGGGGAGAAAGGC 597

QY 555 AGAAGAGATGATCGGTG 571
 DB 598 AGAGAGGATGATGGGAG 614

RESULT 11
 LOCUS PPI489604 1085 bp mRNA linear PLN 11-OCT-2003
 DEFINITION Pinus pinaster mRNA for PII-like protein (glb gene).
 ACCESSION AJ489604
 VERSION AJ489604.1 GI:37653226
 KEYWORDS glb gene; PII-like protein.
 SOURCE Pinus pinaster
 ORGANISM Pinus pinaster
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.

REFERENCE 1
 Cantón, F.J.
 A PII-like protein from xylem of adult Maritime pine trees
 Unpublished
 2 (bases 1 to 1085)
 Cantón, F.J.
 Direct Submission
 Submitted (11-JUN-2002) Cantón F.J., Molecular Biology and Biochemistry, Faculty of Sciences, Campus de Teatinos s/n, E-29071, SPAIN

FEATURES
 Location/Qualifiers
 1..1085
 /organism="Pinus pinaster"
 /mol_type="mRNA"
 /db_xref="taxon:71647"
 /tissue_type="xylem"
 /dev_stage="adult tree"
 1..1085
 /gene="glb"
 42..773
 /gene="glb"
 /function="putative nitrogen sensor protein"
 /codon_start=1
 /product="PII-like protein"
 /protein_id="CA033967.1"

/db_xref="GI:37653227"
 /translation="MAAHLPLVKGSIPLSPSTSSCTSSSISSHTAASPQHFSSP
 AAANKSVKLGTMWRMDGVRSVKHMKQOI KASQAQAPNGTTHPQYVPEA
 NPYKVEAILRPRIRSHVTITGLKMGIRGVTVDVDFGFGVQAGSARQAGSERSKDNFV
 SKVMEIVVSKDQVEAVIDAIIDEARITGIGDKIFVVPADVIRVTRGERDLKPERM
 AGGDQRYLQVYIKRLQATKSLAL"

ORIGIN

Query Match 35.6%; Score 209.6; DB 8; Length 1085;
 Best Local Similarity 70.9%; Pred. No. 4.3e-42;
 Matches 278; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 180 TAGTGCCCAAAATATCTCTGATTATATCCAGACTCGAAATTTTACAAGGTGGAAGCAAT 239
 DB 311 TGGGACCAACATCATCCAGATTATGTCACAGAGCCAACTTTTACAAGTAGAAGCAAT 370

QY 240 TGTGAGACCATGAGAAATCCAGCAAGTTTCATCGGCTTTACTGAAATCGGATTCGAG 299
 DB 371 AITGAGGCCATGGCGCATCTCCCATGTGACTACGGGTCTATTGAAAATGGGATTCGTGG 430

QY 300 TGTACTGTTTCTGATGTGAGAGGTTTGTGTCACAAAGGAGGTTCTACCCAGAGACACGG 359
 DB 431 CGTAACGTCTCTGATGTAGAGGTTTGGAGTTTCAGGCTGGATCTGCACAAACGGCAGC 490

QY 360 TGGCTCTGAGTTCCTCGAAGACAAATTTTGTCTAAAGTTTAAAGATCGAATTCGTTGTA 419
 DB 491 AGCTCCGAGTTTCTTAAGACAAATTTGTGTCAAAAGTAAAGATGAGATTGTGGTATC 550

QY 420 GAAAGACCAAGTGGATTCGTAATCAACACATTAATTAAGGAGCAAGGACAGGAGAT 479
 DB 551 TAAAGATCAGTGAAGACAGTAAATTGATGCAATCATTTGATGAGGCAAGAACTGGAGAAAT 610

QY 480 TGGTGTGTCAGCAATTTTGTGCTGTGTCAGATGTCATAGAGTTAGGACAGGTGA 539
 DB 611 TGGAGATGGCAAAATATTTGTGTTCCAGTTGAGATGTCATCTCGTGTGACACAGGTGA 670

QY 540 GCCTGGGGGAGAAAGACAGAGAGATGATCGTGGT 571
 DB 671 GCGGACTTGAAGCCAGAGAGAAATGCGCTGGT 702

RESULT 12
 LOCUS AC007138 120185 bp DNA linear PLN 01-APR-1999
 DEFINITION Arabidopsis thaliana BAC T7B11 from chromosome IV near 10 cM, complete sequence.
 ACCESSION AC007138
 VERSION AC007138.1 GI:4510323
 KEYWORDS HTG.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 120185)
 Huang, E.N., Nascimento, L., de la Bastide, M., Habermann, K., Vil, M.D., Preston, R.R., Spiegel, L.A., See, L.H., Shah, R., Matero, A., O'Shaughnessy, A., Rodriguez, M., Shekher, M., Swaby, I., Schutz, K., Parnell, L.D., Dedhia, N.N. and McCombie, W.R.
 Arabidopsis thaliana BAC T7B11 from chromosome IV near 10 cM
 Unpublished
 2 (bases 1 to 120185)
 Huang, E.N., Nascimento, L., de la Bastide, M., Habermann, K., Vil, M.D., Preston, R.R., Spiegel, L.A., See, L.H., Shah, R., Matero, A., O'Shaughnessy, A., Rodriguez, M., Shekher, M., Swaby, I., Schutz, K., Parnell, L.D., Dedhia, N.N. and McCombie, W.R.
 Direct Submission
 Submitted (23-MAR-1999) Lita Annenberg Hazen Genome Sequencing Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724
 3 (bases 1 to 120185)
 Parnell, L.D.
 Direct Submission
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE

Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,
Hara, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
Kawai, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Yoshino, M. and Hayashizaki, Y.

TITLE Collection, mapping, and annotation of over 28,000 cDNA clones from
Japanese rice

JOURNAL Science 301 (5631), 376-379 (2003)

MEDLINE 22752273

PUBMED 12869764

REFERENCE 2 (bases 1 to 902)

AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, M.,
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirao, T.,
Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kuribara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K.,
Numasaki, R., Ohneda, E., Ohno, M., Ohtsuka, K., Oka, M., Ooka, H.,
Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Sato, K., Shibata, K.,
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yanagi, W.,
Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and
Yoshimura, A.

Direct Submission

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki
305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,
Tel: 81-29-838-7007, Fax: 81-29-838-7007)

COMMENT This clone is one of the 28K full-length cDNA clones from japonica
rice.

URL : <http://cdna01.dna.affrc.go.jp/cdnap/>
NIRAS Rice Full-length cDNA Project Team: Kikuchi, S., Sato, K.,
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,
Ikeda, R., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,
Ohneda, E., Yanagi, W., Suzuki, K., Li, C., Ohtsuka, K., Shishiki, T. and
Yamamoto, M.
FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y.,
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,
Kodama, T., Kurosaki, R., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,
Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,
Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,
Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kuribara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,
Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,
Yasunishi, A. and Hayashizaki, Y.

FEATURES Location/Qualifiers

source 1..902

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J013149B08"

ORIGIN

Query Match 33.9%; Score 199.4; DB 8; Length 902;
Best Local Similarity 71.3%; Pred. No. 1.6e-39;
Matches 263; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

Qy	208	CGAGCTCGAAATTTACAGGTGGAAGCAATTCGACATCCATGAGCAATCCAGCAAGTT	267
Db	283	CCGGAGTCGGAGTTCTACAGGTGGAAGCAATTCGAGGGCCATCGAGGGTCCCTATGTG	342
Qy	268	TCATCGGCTTTACTGAAATCGGATTCGAGGTCTTACTGTTTCTGATGTGAGGGTTT	327
Db	343	TCATCGGTTTCTGCAATTCGGGATTCGAGGCTGCGGTGTCGAGCTCGGGTTTC	402
Qy	328	GGTGCCACAGAGGAGTTCTTACCGAGAGACACGGTCTGAGTCTTCGGAAGACAAATTT	387
Db	403	GGGCGACAGCGCGGTCAACTGACAGGCATCAAGGGTCAGAAATTCGAGAAGATACATTT	462
Qy	388	GTTCCTAAAGTTAAGTGGAAATCGTTCTTAAAGAGACCAAGTGGAAATCTGTAATCAAC	447
Db	463	ATTGATAAAGTTAAGTGGAAATAGTGGTGTCCAAAGGATCAGGTGGAAGCTGTTGTGAC	522
Qy	448	ACAATAATTGAAGCAGCAAGACAGGAGAGATTGGTGAATGCGCAAGATTTTGTTCCT	507
Db	523	AGATTAATTGAAGCGCCAGACAGGAGAAATTTGGTATGAAATAATTTTGTATACCC	582
Qy	508	GTGTCAGATGTCATAAGAGTTAGGACAGGTGAGCGTGGGAGAAAGACGAGAGATGACT	567
Db	583	GTGTGCGAGCGTGTATGAAATACGACCCGCGAGCGGAGCGGAGGAGGATGGCC	642
Qy	568	GGTGTATG 576	
Db	643	GGAGGGCTG 651	

RESULT 15

AK099152

LOCUS

DEFINITION

AK099152

ACCESSION

AK099152.1

KEYWORDS

FLJ_CDNA; CAP trapper.

SOURCE

Oryza sativa (japonica cultivar-group)

ORGANISM

Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Magnoliophyta; Liliopsida; Poales; Poaceae;

Spermatophyta; Eriophytidae; Erythraeae; Oryza.

REFERENCE

1

AUTHORS

The Rice Full-length cDNA Consortium, National Institute of

Agrobiological Sciences Rice Full-length cDNA Project Team;

Kikuchi, S., Sato, K., Nagata, T., Kawagashira, N., Doi, K.,

Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,

Kojima, K., Namiki, T., Ohneda, E., Yanagi, W., Suzuki, K., Li, C.,

Ohtsuka, K., Shishiki, T., Foundation of Advancement of International

Science Genome Sequencing & Analysis Group; Otomo, Y.,

Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,

Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,

Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,

Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,

Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,

Kawai, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,

Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,

Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,

Yoshino, M. and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from

japonica rice

Science 301 (5631), 376-379 (2003)

JOURNAL

Db	283	CCGAGTCGGAGTTCTACAAAGGTGGAGGCAATCTCGAGGCCATCGAGGGTGCCTTATGTG	342
Qy	268	TCATCGGCTTTACTGAAATCGGATTCGAGGTGTCTACTGTTCTGATGTGAGAGGGTTT	327
Db	343	TCATCGGCTTTGCTGCAATGGGATCGAGGCGTGCAGGTGTCGACGTGCGGGGTTTC	402
Qy	328	GGTGCACAAGGAGGTTCTTACCGAGAGACACGGTGGCTCTGAGTTCTCGGAAGACAAATTT	387
Db	403	GGCGCACAGGCGGGTCAACTGAGAGGCATCAAGGGTCAGAATTTGCGAGAGATACATTT	462
Qy	388	GTTCGTTAAAGTTAAGTGGAAATCGTTGTTAAGAAAGCCAGTGGAAATCTGTAATCAAC	447
Db	463	ATTGATTAAGTTAAGTGGAAATAGTGGTGTCCAGGATCAGGTTGAAAGCTGTTGTTGAC	522
Qy	448	ACAATAATTGAAGGAGCAAGCAGAGAGATTTGGTGTATGGCAAGATTTTGTGTTGCT	507
Db	523	AAGATTAATTGAAGGCAAGCAGAGAAATTTGGTGTATGGAAATATTTTGTATACC	582
Qy	508	GTGTCAGATGTCATAAGAGTTAGGACAGGTGAGCGTGGGGAGAAAGCAGAGAGATGACT	567
Db	583	GTGTCGAGCTGTATCAGAATACCGACCGGAAACGAGGGGAGCGAGCGGAGGATGGCC	642
Qy	568	GGTGATATG 576	
Db	643	GGAGGGCTG 651	

Search completed: May 27, 2004, 22:22:44
Job time : 1740.26 secs

MEDLINE	22752273	
PUBMED	12869764	
REFERENCE	2 (bases 1 to 902)	
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Inamura, K., Itoh, M., Kagawa, I., Kikuchi, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kawama, M., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawasumi, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusuregi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Nishikawa, R., Nishikawa, J., Nishikawa, K., Oka, M., Ooka, H., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Sakai, C., Sakai, K., Osato, N., Ota, Y., Otsu, Y., Ryu, R., Saitoh, H., Saitoh, K., Shibata, K., Sakazume, N., Sato, H., Sasaki, D., Sato, K., Satoh, Y., Sugano, S., Shiragawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, M., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.	
TITLE	Direct Submission	
JOURNAL	Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)	

This clone is one of the 28K full-length cDNA clones from japonica rice.
URL : <http://cdna01.dna.affrc.go.jp/cDNA/>
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Sato, K., Nagata, T., Kawagashira, M., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, M., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otsu, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nishikawa, R., Nishikawa, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imoto, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishikawa, K., Nomura, K., Numasaki, R., Ohno, M., Oota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers
1..902
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="J023070H02"

Query Match 33.9%; Score 199.4; DB 8; Length 902;
Best Local Similarity 71.3%; Pred. No. 1.6e-39;
Matches 263; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

Qy	208	CCAGACTCGAATTTTACAAAGGTGGAGGCAATTCAGACCATCGAGATCCAGCAAGTT	267
----	-----	---	-----

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2004, 19:07:55 ; Search time 1362.19 Seconds

(without alignments)
12890.278 Million cell updates/sec

Title: US-09-756-541-15

Perfect score: 588

Sequence: 1 AAGCGGGCTCAATACGAA.....GTGATATGCTTCACGCT 588

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pin:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_pbg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	481	81.8	549	14	CB261070
2	474	80.6	566	9	AU236084
3	443.8	75.5	685	14	CD825321
4	442.2	75.2	669	14	CD822497

5	442.2	75.2	732	14	CD820903
6	440.6	74.9	745	14	CD819046
7	270.6	46.0	782	14	CD289081
8	262.8	43.0	723	13	BQ798001
9	252.8	43.0	732	9	AJ558383
10	246.6	41.9	597	10	BE239550
11	246.6	41.9	714	12	B1311072
12	246.6	41.9	812	14	CA922829
13	245	41.7	613	9	AM035791
14	245	41.7	702	12	B1932123
15	245	41.7	711	12	B1932336
16	245	41.7	736	12	B1932913
17	245	41.7	741	10	AW738071
18	245	41.7	804	12	B1930336
19	245	41.7	810	12	B1930060
20	244.8	41.6	612	10	BF645522
21	244	41.5	648	9	AI773079
22	243.4	41.4	630	10	AW929534
23	243.4	41.4	685	12	B1929601
24	241.8	41.1	799	12	BM113473
25	240.8	41.0	677	10	BE823431
26	240.2	40.9	756	12	BG595705
27	240.2	40.9	759	14	CF920467
28	239.4	40.7	458	10	BE210153
29	231.8	39.4	560	14	CA514765
30	231.8	39.4	579	10	AW223643
31	224.8	38.2	675	14	CB346213
32	223.2	38.0	719	12	BJ573686
33	222.6	37.9	866	13	C81819
34	222.4	37.8	540	14	CB349912
35	221.8	37.7	551	14	CB349990
36	221.8	37.7	598	14	CB350294
37	221.6	37.7	657	14	CB349285
38	221.6	37.7	701	14	CB349370
39	219.2	37.3	540	10	AW624562
40	218.6	37.2	532	14	CB350222
41	213	36.2	564	13	BQ118918
42	213	36.2	729	14	CF475242
43	212.8	36.2	796	14	CF395143
44	211.8	36.0	673	14	CB085731
45	210.8	35.9	713	13	BX248909

ALIGNMENTS

RESULT 1
CB261070
LOCUS
DEFINITION
33-B9570-013-004-A09-T7R MP1Z-ADIS-012 Arabidopsis thaliana cDNA
clone MP1Zp769A094Q 5-PRIME, mRNA sequence.
549 bp mRNA linear EST 06-NOV-2003

ACCESSION
CB261070
VERSION
CB261070.1
KEYWORDS
GI:32885843

SOURCE

ORGANISM
Arabidopsis thaliana (thale cress)

REFERENCE
1 (bases 1 to 549)

AUTHORS
Schmid, K.J., Soerensen, T.R., Stracke, R., Torjek, O., Altmann, T., Mitchell-Olds, T. and Weisshaar, B.

TITLE
Large-scale identification and analysis of genome-wide

JOURNAL
Genome Res. 13 (6), 1250-1257 (2003)

MEDLINE
22683290

PUBMED
12799357

COMMENT
Contact: Weisshaar B

ADIS DNA core facility at MP1Z

Max-Planck-Institute for Plant Breeding Research

Carl-von-Linne Weg 10, 50829 Koeln, Germany

Fax: 00492215062851

Email: weisshaar@mpiz-koeln.mpg.de

CD820903 BN25.040D
CD819046 BN20.047L
CD289081 V-B-112F0
BQ798001 EST 6939
AJ558383 AJ558383
BE239550 EST403599
B1311072 EST531282
CA922829 EST640547
AW035791 EST281945
BI932123 EST552012
BI932336 EST552225
BI932913 EST552802
BI938071 EST339498
BI930336 EST550225
BI930060 EST549949
BF645522 NF036F08E
AI773079 EST254179
AW929534 EST338322
BI929601 EST549490
BM113473 EST561009
BE823431 GM700019B
BG595705 EST494383
CF920467 gmrhww3-
BE210153 S039003.Y
CA514765 K509030G0
AW223643 EST300454
CB346213 CAB2SG000
BJ573686 BJ573686
C81819 C81819 Citr
CB349912 CAB2SG000
CB349990 CAB2SG000
CB350294 CAB2SG000
CB349285 CAB2SG000
CB349370 CAB2SG000
AW624562 EST322507
CB350222 CAB2SG000
BQ118918 EST604494
CF475242 RTW2_14
CF395143 RTD2_9_5
CB085731 hgl509_9
BX248909 BX248909

Insert Length: 549 Std Error: 0.00
 Plate: 4 Row: A Column: 09
 Seq primer: T7R; CTAATACGACTCACTATAGGA.
 Location/Qualifiers

FEATURES

source

1..549
 /organism="Arabidopsis thaliana"
 /mol_type="mrna"
 /cultivar="Landsberg erecta (Ler)"
 /db_xref="GABI:589834"
 /db_xref="taxon:3702"
 /clone="MPIZp763A09C"
 /tissue_type="whole plant"
 /dev_stages="adult plant, mixed stresses"
 /lab_host="E. coli XL1-Blue MRF"
 /clone_lib="MPIZ-ADIS-012"
 /note="Vector: pSPORT1; Site 1: Sali; Site 2: NotI; cDNA library from Arabidopsis thaliana, accession Landsberg erecta; six weeks old total plants grown under long-day conditions in soil, whole adult plants were treated for 24 hours with different stresses, (1) at 4M-0 C in the dark, (2), at 37 Grad C in the dark, (3) lying in the lab after removing from soil, (4) in the greenhouse after wounding leaves with a forceps, (5) in the lab watering with a 150 mM NaCl solution, (6) at 26 Grad C in the light/UV; equal quantities of stressed plant material were pooled; library was made at the Max-Planck-Institute for Plant Breeding Research, Cologne, Germany; cloning sites Sali-NotI, primer sites and orientation:
 T7-Sali-CACGGGTGCG-5prime-cDNA-polyA-CC-NotI-SP6; Note: Sequencing granted in the context of the GABI Arabidopsis Verbund I: Genetic Diversity, 'Establishment of high-efficiency SNP-based mapping tools and development of methods for genome-wide mutation detection' Pi: Bernd Weisshaar Sequence submission managed by RZPD/GABI-Primary database: http://gabi.rzpd.de. This clone is available from RZPD; contact RZPD (clone@rzpd.de) for further information."

ORIGIN

Query Match 81.8%; Score 481; DB 14; Length 549;
 Best Local Similarity 97.7%; Pred. No. 1e-124;
 Matches 509; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

Cy 1 ATGGGGGGTCAATGACGAACCCATCTCAATACTCTCTCGGTTCTATCTCATCGA 60
 Db 30 ATGGAGGGGTCAATGACGAACCCATCTCAATACTCTCTCGGTTCTATCTCATCGA 89
 Cy 61 AAGAACATTTGCTTTCTCTGATTGCAATTCGATTGTTCTGATTGATTCCCGACCA 120
 Db 90 AAG-ACATGCTTTCTCTGATTGCAATTCGATTGTTCTGATTGATTCCCGACCA 148
 Cy 121 TCTTGGCTCGATTGGTCAAAAGTCACGAGTAATAACAGTCGTTGTTTACCTGCTGTT 180
 Db 149 TCTTGGCTCGATTGGTCAAAAGTCACGAGTAATAACAGTCGTTGTTTACCTGCTGTT 208
 Cy 181 AGTGCCCAATATCTCTGATTATATCCAGACTC-GAAATTTTCAAGTGGAGGACAT 239
 Db 209 AGAGCCCAAGCTCTCTGATTATATCCAGACTCAGAAATTTTCAAGTGGAGGACAT 268
 Cy 240 TGTGAGCATCGAGATCCAGAGTTTCTATCGGCTTTTACTGAAATCGGATTCGAGG 299
 Db 269 TGTGAGCATCGAGATCCAGAGTTTCTATCGGCTTTTACTGAAATCGGATTCGAGG 328
 Cy 300 TGTACTGTTTCTGATGTCAGAGGGTGTGGTGCACAGGAGTTTCTACCGAGACACGG 359
 Db 329 TGTACTGTTTCTGATGTTCCAGGGGTTTGGTGCACAGGAGGTTTCTACTGAGACACGG 388
 Cy 360 TGGCTCTGAGTTCTCGGAGACAAATTTGTTCTAAAGTTTACATGGAATCGTTGTTAA 419
 Db 389 TGGCTCTGAGTTCTCGGAGACAAATTTGTTCTAAAGTTTACATGGAATCGTTGTTAA 448
 Cy 420 GAAAGACCAAGTGGATCTGTAATCAACAATAATTGAAGGAGCAAGGACGAGGAGAT 479

Db 449 GAAAGACCAAGTGGATCTGTAATCAACAATAATTGATGGAGCAGGACGAGAGAT 508
 Cy 480 TGGTATGGCGAAGATTTTGTGTTTGGCTGTGTGATGATGCA 520
 Db 509 TGGTATGGCGAAGATTTTGTGTTTGGCTGTGTGATGATGCA 549

RESULT 2

AU236084

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

ORGANISM

Arabidopsis thaliana (Chale cress)

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 566)

Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,

Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,

Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.

and Shinozaki, K.

Large scale analysis of Arabidopsis full-length cDNA (2002b)

Unpublished (2002)

Contact: Motoaki Seki

Plant Functional Genomics Research Group

RIKEN Genomic Sciences Center

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

Tel: 81-298-36-4359

Fax: 81-298-36-9060

Email: msekic@riken.go.jp

An Arabidopsis full-length cDNA library was constructed essentially

as reported previously (Seki et al., 1998). cDNA cleaved with BamHI

and XhoI was ligated to modified lambda FLC-1 vector (Carninci et

al., submitted for publication) digested with BamHI and Sali. This

clone is in a modified pBluescript vector. Please visit our web

site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further

details.

Location/Qualifiers

1..566

/organism="Arabidopsis thaliana"

/mol_type="mrna"

/db_xref="taxon:3702"

/clone="RAF14-63-K15"

/tissue_type="root"

/lab_host="DH10B"

/clone_lib="RAF14"

/note="Site_1: BamHI; Site_2: Sali"

ORIGIN

Query Match 80.6%; Score 474; DB 9; Length 566;
 Best Local Similarity 97.8%; Pred. No. 9.9e-123;
 Matches 500; Conservative 0; Mismatches 8; Indels 3; Gaps 2;

Cy 3 GCGCGGTCAATGACGAACCCATCTCAATACTCTCTCGGTTCTATCTCATCGAAA 62
 Db 55 GCGCGGTCAATGACGAACCCATCTCAATACTCTCTCGGTTCTATCTCATCGAAA 114
 Cy 63 GAACATGCTTTCTCTGATTGCAATTCGATTGTTTCTGATTGATTCGACCATC 122
 Db 115 GAACATGCTTTCTCTGATTGCAATTCGATTGTTTCTGATTGATTCGACCATC 174
 Cy 123 TTGCTCGATTGGTGCACAAAGTCACCGAGTAATAACAGTCGTTGTTTACCTGCTTAG 182
 Db 175 TTGCTCGATTGGTGCACAAAGTCACCGAGTAATAACAGTCGTTGTTTACCTGCTTAG 234
 Cy 183 TGGCCAAATATCTTCTGATTATATCCAGCTCGAAATTTTCAAGTCGAGCAATTGT 242
 Db 235 TGGCCAAATATCTTCTGATTATATCCAGCTCGAAATTTTCAAGTCGAGCAATTGT 294

```

QY 243 CAGACCATGGAGAAATCCAGCAAGTTTTCATCGGCTTTACTGAAATTCGGGATTCGAGGTGT 302
DB |||||
DB 295 CAGACCATGGAGAAATCCAGCAAGTTTTCATCGGCTTTACTGAAATTCGGGATTCGAGGTGT 354
QY 303 TACTGTCTTCTGATGTCAGAGGGTTTGGTGCACAAGGAGGTTCCTACCGAGAGACACGGTGG 362
DB |||||
DB 355 TACTGTCTTCTGATGTCAGAGGGTTTGGTGCACAAGGAGGTTCCTACCGAGAGACACGGTGG 414
QY 363 CTCTGAGTTCTCGAAGACAAATTTTGTCTGAAGTTAAGATGGAAATCGTTGTTAAGAA 422
DB |||||
DB 415 CTCTGAGTTCTCGAAGACAAATTTTGTCTGAAGTTAAGATGGAAATCGTTGTTAAGAA 474
QY 423 AGACCAAGT--GGAATCTGTAATCAACACAAATTAATCAAGGAGCAAGG--ACAGGAGAGAT 479
DB |||||
DB 475 AGACCAAGTGGGAATCTGTAATCAACAAATTAATCAAGGAGCCAGGACAGGAGAGAT 534
QY 480 TGGTGATGGCAAGATTTTGTGCTGTG 510
DB |||||
DB 535 TGGTGATGGCAAGATTTTGTGCTGTG 565

CD825321 685 bp mRNA linear EST 10-JUL-2003
LOCUS BN25_060G17FC11129 BN25 Brassica napus cDNA clone BN25060G17, mRNA
DEFINITION sequence.
ACCESSION CD825321
VERSION CD825321.1 GI:32507261
KEYWORDS EST.
SOURCE Brassica napus (rape)
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 685)
AUTHORS Genoplante.
TITLE Genoplante, a major partnership french program in plant genomics
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (http://www.genoplante.com
and http://genoplante-info.infobiogen.fr).
FEATURES
Location/Qualifiers
source
1..685
/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="J6t neuf"
/db_xref="taxon:3708"
/clone="BN25060G17"
/tissue_type="seed"
/clone_lib="BN25"

ORIGIN
Query Match 75.5%; Score 443.8; DB 14; Length 685;
Best Local Similarity 85.6%; Pred. No. 3.5e-114;
Matches 506; Conservative 0; Mismatches 82; Indels 3; Gaps 1;

QY 1 ATGGCGCGTCATGACGAACCCATCTCAATACCTCTCTCGGTTCTTATCTGATCGA 60
DB |||||
DB 9 ATGGCGCGTTCAATCGCGAAACCCATCTCTCGAACCTCTCTCGGTTCTTACTCCGACCGA 68
QY 61 AAG---AACATTGCTTTCTGATGTCGATTTGTTCTGGAATCAGACATTCGCGA 117
DB |||||
DB 69 AAGAACACATTTCCCTTCTGATGTCATCTCGATATCTCTGATATCTCTGGTTCCGACATTCGCGA 128
QY 118 CCATCTTGCCTCGATTTGGTCACAAAGTACCGAGTAATACAGTGGTGTGTTTACCTGTC 177
DB |||||
DB 129 CCGTCTTCCCTCGATTTGGTCGCAAAACCCACCAACCGACCAAGGCTCGGTTTCCCATCTC 188

```

```

QY 178 GTTAGTGCCCAATATATCTTCTGATTATATTCAGACTCGAAATTTTACAAAGGTGGAGCA 237
DB |||||
DB 189 GTTAGAGCCCAAGCTCTTATGATTATATTCAGACTCAAAGTTTACAAAGGTGGAGCA 248
QY 238 ATTGTCAAGCCATGGAGAAATCCAGAAAGTTTCAATCGCTTTTACGAAAAATCGGGATTGGA 297
DB |||||
DB 249 ATTGTCAAGCCATGGAGAAATCCAGCAATGTTTTCATCGGCTTTTACTGAAAAATCGGGATTGGA 308
QY 298 GGTGTTACTGTTTCTCATGTGTGAGAGGGTTTGGTGCAACAAGGAGTCTTACCGAGAGACAC 357
DB |||||
DB 309 GGTGTACCGGTTTCTCATGTGTGAGAGGGTTTGGTGCAACAAGTGTGATCCAAAGAGAGACAT 368
QY 358 GGTGGGCTCTGAGTTCTCGGAAGACAAATTTGTGCTAAAGTTTAAGATGGAATCGTTGTT 417
DB |||||
DB 369 GGTGGGCTCTGAGTTCTCTGAAGACAAAGTTGTTGCTAAAGTTTAAGATGAGATCGTTGTT 428
QY 418 AAGAAAGACCAAGTGGAAATCTGTAAATCAACACAAATTAATCAAGGAGCAAGGAGAG 477
DB |||||
DB 429 AAGAAAGACCAAGTGGAGTCTGTAATCAACACCAATTAATCGATGGAGCAAGAACAGGAGAG 488
QY 478 ATTGTTGATGGCAAGATTTTGTGCTGCTGTGTCAAGATGTCAAGAGTTAGGACAGGT 537
DB |||||
DB 489 ATCGGAGATGGCAAAATATTCGTTTTCCTGTTGTGATGTCATGATGTCATGAGGTTTCGACAGGT 548
QY 538 GAGCGTGGGAGAAAGCAGAGAAAGATGACTGGTGATATGCTTCCCGTCT 588
DB |||||
DB 549 GAGCGTGGAGAAACAGCAGAGAGAGATGACTGGTGACATGCTTTCGTCTATCT 599

CD822497 669 bp mRNA linear EST 10-JUL-2003
LOCUS BN25_045G20F020108 BN25 Brassica napus cDNA clone BN25045G20, mRNA
DEFINITION sequence.
ACCESSION CD822497
VERSION CD822497.1 GI:32504437
KEYWORDS EST.
SOURCE Brassica napus (rape)
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 669)
AUTHORS Genoplante.
TITLE Genoplante, a major partnership french program in plant genomics
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (http://www.genoplante.com
and http://genoplante-info.infobiogen.fr).
FEATURES
Location/Qualifiers
source
1..669
/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="J6t neuf"
/db_xref="taxon:3708"
/clone="BN25045G20"
/tissue_type="seed"
/clone_lib="BN25"

ORIGIN
Query Match 75.2%; Score 442.2; DB 14; Length 669;
Best Local Similarity 85.4%; Pred. No. 9.8e-114;
Matches 505; Conservative 0; Mismatches 83; Indels 3; Gaps 1;

QY 1 ATGGCGCGTCATGACGAACCCATCTCAATACCTCTCTCGGTTCTTATCTGATCGA 60
DB |||||
DB 2 ATGGCGCGTTCAATCGCGAAACCCATCTCTCGAACCTCTCTCGGTTTCTACTCCGACCGA 61

```


LOCUS	BQ798001	723 bp	mRNA	linear	EST 30-JUL-2002
DEFINITION	EST 6939 Ripening Grape berries Lambda Zap II Library Vitis vinifera cDNA clone RT093C02 3', mRNA sequence.				
ACCESSION	BQ798001				
VERSION	BQ798001.1	GI:22012967			
KEYWORDS	EST.				
SOURCE	Vitis vinifera				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; Vitaceae; Vitis.				
REFERENCE	1 (bases 1 to 723)				
AUTHORS	Abbal, P., Agasse, A., Ageorges, A., Atanassova, R., Barrieu, F., Couture, C., Dedaldechamp, F., Delrot, S., Glissant, D., Grimplet, J., Hamdi, S., Romieu, C. and Terrier, N.				
TITLE	Generation of Expressed Sequence Tag from Grape Berry (skin, pulp or seeds) at Various Developmental Stages				
JOURNAL	Unpublished (2002)				
COMMENT	Contact: Romieu C. Unite de Recherche des Produits de la Vigne Institut National de la Recherche Agronomique 2, place Viala, 34 060 Montpellier Cedex 01, France Tel: 00-33-(0)4-99-61-28-62 Fax: 00-33-(0)4-99-61-28-57 Email: romieu@ensam.inra.fr Seq primer: T7.				
FEATURES	Location/Qualifiers 1..723 /organism="Vitis vinifera" /mol_type="mRNA" /cultivar="Shiraz" /db_xref="taxon:29760" /clone="RT093C02" /dev_stage="ripening stage" /clone_lib="Ripening Grape berries Lambda Zap II Library" /note="Organ: Fruit; Vector: Lambda Zap II; Site 1: Eco RI; Site 2: XhoI; Oriented library, construction described in Generation of ESTs from grape Berry (skin, pulp or seeds) at various developmental stages by Terrier, N., Ageorges, A., Abbal, P., Romieu, C. in J. Plant Physiol. 158 (12): 1575-83 2001"				
ORIGIN					
Query Match	44.7%;	Score 262.8;	DB 13;	Length 723;	
Best Local Similarity	76.8%;	Pred. No. 4.4e-63;			
Matches 321;	Conservative	0;	Mismatches 97;	Indels	0; Gaps 0;
Qy	154	AATAACAGTCGTGTTTACCTGCTGCTAGTCCCAATATCTTCTGATTATATCCAGAC			
Db	675	AAGATGGATCTGTTCTTCCCATTTGTCCAGGCTCAGAGCTCTCCAGTTATACCCAGAC			
Qy	214	TCGAATTTTCAAGTGGAGCAATTTGTCCAGCATCCAGATCCAGAGTTTTCATCG			
Db	615	GCGCAGTTTTCACAAAGTTGAAGCGATCTCGAGGCCCTGCGCAATCCAGCAGTTTCTTCG			
Qy	274	GCTTTACTGAAATCGGATTCGAGGTGTTACTGTTTCTGATGTGAGAGGTTTGGTGCA			
Db	555	GCTTTGCTGAAATGGGTATTCGTGGTGTACTGTTTCTGATGTCCGGGGCTTTGGTGCT			
Qy	334	CAAGGAGTTTCTACCGAGAGACCGTGGCTCTGAGTTTCTCGGAAGACAAATTTGGTCT			
Db	495	CAAGCGGTTCCGCTGAAAGACAGGCGCTCTGAAATTTCTGAGGACAAATTTGGTCT			
Qy	394	AAAGTTAAGATGGAATTCGTTTGAAGAAGACCAAGTGGAAATCTGTATCAACACATA			
Db	435	AAAGTCMAATGGAGATTGTTGGTAGCAAGACCCAGGTTGAAGCAGTAATTCACAATC			
Qy	454	ATTGAAGGACCAAGCAGGAGAGATTCGTGATGCGAAGATTTTGTGCTGTGTC			
Db	375	AATGAGTTGGCAGGACTGGAGAGATTGGTGTGATGCGAAGATCTTTTGTGTCCTATCG			
Qy	514	GATGTCATAAGAGTTAGGACAGCTGAGCGTGGGAGAAAGGCTGAGAGATGACTGGTG			

LOCUS	AJ558383	732 bp	mRNA	linear	EST 12-JUN-2003
DEFINITION	AJ558383 Antirrhinum majus whole plant Antirrhinum majus cDNA clone 018.1.07.f22, mRNA sequence.				
ACCESSION	AJ558383				
VERSION	AJ558383.1	GI:31660955			
KEYWORDS	EST.				
SOURCE	Antirrhinum majus (snapdragon)				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamials; Lamiales; Antirrhinaceae; Antirrhineae; Antirrhinum.				
REFERENCE	1 (bases 1 to 732)				
AUTHORS	Zachgo, S., Stueber, K., Saedler, H., Sommer, H. and Schwarz-Sommer, Z.				
TITLE	Antirrhinum EST collection				
JOURNAL	Unpublished (2003)				
COMMENT	Contact: Schwarz-Sommer Z Molekulare Pflanzen-genetik MPI fuer Zuechtungs-forschung Carl-von-Linne Weg 10, D-50829, Germany.				
FEATURES	Location/Qualifiers 1..732 /organism="Antirrhinum majus" /mol_type="mRNA" /db_xref="taxon:4151" /clone="018.1.07.f22" /tissue_type="whole plant" /clone_lib="Antirrhinum majus whole plant"				
ORIGIN					
Query Match	43.0%;	Score 252.8;	DB 9;	Length 732;	
Best Local Similarity	77.0%;	Pred. No. 2.9e-60;			
Matches 308;	Conservative	0;	Mismatches 92;	Indels	0; Gaps 0;
Qy	172	CTGTGCTGTAGTCCCAAAATATCTTCTGATTATATTCAGACTCGAATTTTACAGGTG			
Db	172	CCAATTGTTAGGGCTCAGATTCTCCAGAAATCCCTACCTGACTCCAAAGTTCTACAAAGTC			
Qy	232	GAAGCAATTTGTACAGACCATGGAGATCCAGCAAGTTTTCATCGGCTTTACTGAAATCGGG			
Db	232	GAAGCTATTTCTAAGACCCCTGGAGGATCCACAGGTTTCTCGGCCCTGCTGAAATGGGA			
Qy	292	ATTCGAGGTGTTTACTGTTTCTGATGTGAGAGGTTTGGTGCACAGAGGTTTCTACCGAG			
Db	292	ATTCGTGGTGTCTACTGTCTCTGATGTAOGAGGGTTTGGTTCTCAGGGTGGTTTGACAGAA			
Qy	352	AGACACGGTGGCTCTGAGTTCTCGGAAGCAAAATTTGTTGCTTAAAGTTAAGATGGAATC			
Db	352	AGACAGTTTGGCTCTGAAATTTCTCAGAAGCAAAATTTGTTGCAAAAGTTAAGATGGAGATT			
Qy	412	GTTTGTTAAGAAAGACCAAGTGGAAATCTGTAATCAACAATAATTAAGAGAGCAGAGACA			
Db	412	GTTGTGACCAAGACCAAGGTTGATGCGGTAAATAGACAAGTAATAGAGGACCAAGACA			
Qy	472	GGAGAGATTGGTGTGTCAGAGATTTTGTGTTTGCCTGTGTGATGTCATAGAGATTAGG			
Db	472	GGAGAGATTGGTGTGTCGAAAGATTTTGTGCTTCTCCAGTTTCTGATGTAATAGAGTTGCG			
Qy	532	ACAGTGTGAGCTGGGGAGAAAGCAGAGATGACTGGTG			
Db	532	ACTGGTGACAGGGGGAAAGGCTGAGAGGATGGTGGGAG			

LOCUS	BE239550	597 bp	mRNA	linear	EST 12-JUL-2000
DEFINITION	BE239550 MHRP- Medicago truncatula cDNA clone pmHRP-2Bf3, mRNA sequence.				

```

ACCESSION      BE239550
VERSION        BE239550.1  GI:9055733
KEYWORDS       EST.
SOURCE         Medicago truncatula (barrel medic)
ORGANISM       Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
               rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
               Medicago.
REFERENCE      1 (bases 1 to 597)
AUTHORS       Harrison,M.J., Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S.,
               Holt,I.E., Cho,J. and Fraser,C.M.
TITLE         ESTs from phosphate-starved roots of Medicago truncatula
JOURNAL       Unpublished (2000)
COMMENT       Contact: Maria J. Harrison
               Plant Biology Division
               The Samuel Roberts Noble Foundation
               2510 Sam Noble Parkway, Ardmore, OK 73401, USA
               Tel: 580-223-5810
               Fax: 580-221-7380
               Email: mjharrison@noble.org
               The Samuel Noble Roberts Foundation: N265417e
               TIGR sequence name: MTHAC267K
               More information is available at:
               http://chryslie.tamu.edu/medicago
               Seq primer: SKmod (CTA GAA CTA gtg gAT CC).
               Location/Qualifiers
               1..597
               /organism="Medicago truncatula"
               /mol_type="mRNA"
               /cultivar="A17"
               /db_xref="taxon:3880"
               /clone="PMERP-28F3"
               /tissue_type="roots"
               /dev_stage="phosphate-starved"
               /lab_host="XILOR"
               /clone_lib="MHRP"
               /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
               XhoI; At the trifoliolate stage, M. truncatula plants were
               transplanted to phosphate-free sand and grown for a
               further 30 days. During this period, they were fertilized
               twice weekly with 1/2 Hoaglands solutions containing 20uM
               potassium phosphate. cDNA was prepared from polyA+
               enriched RNA. The cDNA was directionally ligated into the
               Unizap Xa vector from Stratagene and packaged using
               Gigapack III Gold packaging extracts. Plasmids containing
               cDNA inserts were excised from the recombinant lambda-Zap
               phage using Ex-assist helper phage and propagated in
               XILOR cells."

FEATURES             source
     source
     1..714
     /organism="Medicago truncatula"
     /mol_type="mRNA"
     /cultivar="A17"
     /db_xref="taxon:3880"
     /clone="PGESD9J4"
     /tissue_type="immature seeds"
     /dev_stage="Immature seeds, 11 to 19 days after
     pollination"
     /clone_lib="GESD"
     /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
     XhoI; Immature seeds, collected from pods ranging in age
     from 11 to 19 days after pollination, were harvested from
     greenhouse-grown plants. Seeds were removed and
     separated from pod walls and were immediately frozen in
     liquid nitrogen. Seeds throughout the age range were
     pooled for mRNA extraction. cDNA was prepared from polyA+
     enriched RNA. The cDNA was directionally ligated into
     the Unizap Xa vector from Stratagene and packaged using
     Gigapack III Gold packaging extracts. Plasmids containing
     cDNA inserts were excised from the recombinant lambda-Zap
     phage using Ex-assist helper phage and propagated in
     XILOR cells."

ORIGIN
Query Match      41.9%; Score 246.6; DB 10; Length 597;
Best Local Similarity 75.6%; Pred. No. 1.6e-58;
Matches 306; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 167 TTTTACCTGCTAGTGGCCAAATCTCTGATTAATTCAGACTCGAAATTTTACA 226
DB 173 TTCTTCCAAATCAGAGCTCAAAACCTTCTGACTATGTTCTGTAATCCAGTTTACA 232
QY 227 AGGTGGAACCAATTGTCAGACCATGAGAAATCCAGCAAGTTTTCATCGGCTTTACTGAAA 286
DB 233 AAGTTGAACCAATCTCAGGCCATGAGAAATCCCTCAGGTTTCTCGGTTTGTGAAA 292
QY 287 TCGGGATTCGAGGTGTTACTGTTTCTGATGTGAGAGGTTTGGTGACAGGAGTTCTA 346
DB 293 TCGGAATTCGTGTCACGTATCTGATGTCAGGCTTTTGGTGTCTCAGGTTGGCTCAA 352
QY 347 CCAGAGACACGTTGCTGAGTTCTCGGAACAAATTTGTTGCTAAAGTTAAGATGG 406
DB 353 AAGAGAGGAGGAGGCTCGGAATTTTCTGAGACAAATTTTGTGTCAAAGTTTAAATGG 412
QY 407 AATTCCTGTTTGAAGAGACCAAGTGGAACTCTGTAATCAACAAATTAATGAAGGACAA 466

```

```

Db 413 AAATAGTGGTGAGAAAACACACAGGTTGAGCCAGTGATATAACAAAATATATGAGACGGCAA 472
QY 467 GGACAGGAGAGATTGGTGATGCGACAGATTTTGTGCTGCTGTGATGATGTCATAGAG 526
DB 473 GAACTGGGGAGATTGGTGATGCGCAAAATTTCTTGATCCCTGTATCTGATGTAATAAGAA 532
QY 527 TTAGCACAGGTCAGGCTGGGAGAAAGACAGAGAGACTGCTGGTG 571
DB 533 TCCGCACAGGTGAGCTGGGAGCAGGCTGAGAGAGATGCTGGGG 577

```

```

RESULT 11
BI311072
LOCUS      714 bp mRNA linear EST 20-JUL-2001
DEFINITION      EST5312822 GESD Medicago truncatula cDNA clone pGESD9J4 5' end,
               mRNA sequence.
ACCESSION      BI311072
VERSION        BI311072.1  GI:14985399
KEYWORDS       EST.
SOURCE         Medicago truncatula (barrel medic)
ORGANISM       Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
               rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
               Medicago.

```

```

REFERENCE      1 (bases 1 to 714)
AUTHORS       Grusak,M.A., Samac,D.A., Town,C.D., Van Aken,S., Utterback,T.,
               Cho,J. and Fraser,C.M.
TITLE         ESTs from developing reproductive tissues of Medicago truncatula
JOURNAL       Unpublished (2001)
COMMENT       Contact: Michael A. Grusak
               USDA/ARS Children's Nutrition Research Center
               Baylor College of Medicine
               1100 Bates Street, Houston, TX 77030-2600, USA
               Tel: 713-798-7044
               Fax: 713-798-7078
               Email: mgrusak@bcm.tmc.edu
               B397927e
               TIGR sequence name: MTPAP50TK
               More information is available at: www.medicago.org
               Seq primer: SKmod (CTA GAA CTA gtg gAT CC).

```

```

FEATURES             Location/Qualifiers
     1..714
     /organism="Medicago truncatula"
     /mol_type="mRNA"
     /cultivar="A17"
     /db_xref="taxon:3880"
     /clone="PGESD9J4"
     /tissue_type="Immature seeds"
     /dev_stage="Immature seeds, 11 to 19 days after
     pollination"
     /clone_lib="GESD"
     /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
     XhoI; Immature seeds, collected from pods ranging in age
     from 11 to 19 days after pollination, were harvested from
     greenhouse-grown plants. Seeds were removed and
     separated from pod walls and were immediately frozen in
     liquid nitrogen. Seeds throughout the age range were
     pooled for mRNA extraction. cDNA was prepared from polyA+
     enriched RNA. The cDNA was directionally ligated into
     the Unizap Xa vector from Stratagene and packaged using
     Gigapack III Gold packaging extracts. Plasmids containing
     cDNA inserts were excised from the recombinant lambda-Zap
     phage using Ex-assist helper phage and propagated in
     XILOR cells."

```

```

ORIGIN
Query Match      41.9%; Score 246.6; DB 12; Length 714;
Best Local Similarity 75.6%; Pred. No. 1.6e-58;
Matches 306; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 167 TTTTACCTGCTAGTGGCCAAATCTCTGATTAATTCAGACTCGAAATTTTACA 226

```

Db 172 TTCTTCCAAATCAGAGCTCAAAACCTTCTGACTATGTCTCGAATCCAAAGTTTACA 231
 QY 227 AGGTGAAGCAATGTGAGCCATCGAGAGATCCAGAGTTTCTCGCTTACTGAAAA 286
 Db 232 AAGTTGAAGCAATCTCAGGCCATCGAGAGATCCCTCAGGTTTCTTCGCGTTTGTGAAA 291
 QY 287 TCGGATTCGAGGTGTTACTGTTTCTGATGTGAGAGGTTTGGTGCACAGGAGCTCTA 346
 Db 292 TGGGAATTCGTGGTGTCACTGTATCTCTGATGTCAAGGGGTTTGGTGCCTCAGGCTGCTCAA 351
 QY 347 CCGAGACACGCGTCTGAGTTCTCGAAGACAAATTTGTTGCTTAAAGTTAAGATGG 406
 Db 352 AAGAGCGCAGGAGGCTCCGAATTTCTGAAGCAATTTTGTGGCCAAAGTTAAATGG 411
 QY 407 AAATCGTTGTTAAGAAAGACCAAGTGAATCTGTAATCAACACAAATTAATGAAGAGCAA 466
 Db 412 AAATAGTGTGAGAAAGACCAAGGTTGAGGAGTGATAAAACAAATATATGAGAGCGGCAA 471
 QY 467 GGACAGAGAGATGTTGATGCGCAAGATTTTGTGTTGCTGCTGTCAGATCTCATAGAG 526
 Db 472 GAATCGGGAGATGGTGATGGCAAAATTTTCTGATCCCTGTATCTGATGTAAATAGAA 531
 QY 527 TTAGCACAGGTGACGTTGGGAGAGAAAGCAGAGAGATGACTGGTG 571
 Db 532 TCCGACAGGTGAGCGTGGGAGGAGCAGGCTGAGAGATGCTGGG 576

RESULT 12
 CA922829/c
 LOCUS
 DEFINITION EST640547 MTUS Medicago truncatula cDNA clone MTUS-58H7, mRNA sequence.
 ACCESSION CA922829
 VERSION CA922829.1 GI:27409759
 KEYWORDS EST.
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
 REFERENCE 1 (bases 1 to 812)
 VandenBosch, K., Endre, G., Silverstein, K., Town, C.D., Van Aken, S., Utterback, I., Cheung, F. and Fraser, C.M.
 The Medicago truncatula 6x unigene set: cDNA clones selected and re-arrayed from various libraries
 Unpublished (2002)
 CONTACT: VandenBosch K
 Department of Plant Biology
 University of Minnesota
 220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
 Tel: 612 624 2755
 Fax: 612 625 1738
 Email: kvandenb@cbs.umn.edu
 Alias Clone pMHRP-28F3
 TIGR sequence name: MTUCF91TV
 More information is available at: www.medicago.org
 Seq primer: (gtA tA cGg cTc Act AtA ggg C).
 Location/Qualifiers
 1..812
 /organism="Medicago truncatula"
 /mol_type="mRNA"
 /cultivar="Al7"
 /db_xref="taxon:3880"
 /clone="MTUS-58H7"
 /tissue_type="mixed tissues"
 /dev_stage="various stages"
 /lab_host="XLOLR"
 /clone_lib="MTUS"

FEATURES
 source
 1..812
 /organism="Medicago truncatula"
 /mol_type="mRNA"
 /cultivar="Al7"
 /db_xref="taxon:3880"
 /clone="MTUS-58H7"
 /tissue_type="mixed tissues"
 /dev_stage="various stages"
 /lab_host="XLOLR"
 /clone_lib="MTUS"
 /note="Vector: pBluescript SK-; Site: 1: EcoRI; Site 2: XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging

extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XLOLR cells."

ORIGIN
 Query Match 41.9%; Score 246.6; DB 14; Length 812;
 Best Local Similarity 75.6%; Pred. No. 1.7e-58;
 Matches 306; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
 QY 167 TTTTACCTGCTGTAGTGCCTCAAAATATCTTCTGATTATATCCAGACTCGAAATTTTACA 226
 Db 623 TTTTCCAAATCAGAGCTCAAAACCTTCTGACTATGTTCCTGAAATCCAGTTTACA 564
 QY 227 AGGTGAAGCAATTTCTGAGACCATGGAGAAATCCAGCAAGTTTCACTGCAAAA 286
 Db 563 AAGTTGAAGCCATTTCTAGGCCATGGAGAAATCCCTCAGTTTCTTCGGGTTTGTGAAA 504
 QY 287 TCGGATTCGAGGTGTTACTGTTTCTGATGTGAGAGGTTTGGTGCACAGAGGTTCTA 346
 Db 503 TGGGAATTCGTGGTGTCACTGTATCTGATGTCAAGGGGTTTGGTCTCAGGGTGGCTCAA 444
 QY 347 CCGAGACACGTTGGCTCTGAGTTCTCGAGACAAATTTGTTGCTTAAAGTTAAGATGG 406
 Db 443 AAGAGAGCGCAGGAGGCTCCGAATTTTCTGAAGCAATTTTGTTCCTCAAGTTAAATGG 384
 QY 407 AAATCGTTGTTAAGAAAGACCAAGTGAATCTGTAATCAACACAAATTAATGAAGAGCAA 466
 Db 383 AAATAGTGTGAGAAAGACCAAGTTGAGGCAATTAACAAATTAATGAAGAGCGCAA 324
 QY 467 GGACAGAGAGATGTTGATGCGCAAGATTTTGTGTTGCTGCTGTCAGATCTCATAGAG 526
 Db 323 GAATCGGGAGATGGTGTGATGCGCAAAATTTTCTGATCCCTGTATCTGATGTAAATAGAA 264
 QY 527 TTAGCACAGGTGAGCGTGGGAGAGAAAGCAGAGAGATGACTGGTG 571
 Db 263 TCCGACAGGTGAGCGTGGGAGGAGGCTGAGAGATGCTGGG 219

RESULT 13
 AM035791
 LOCUS
 DEFINITION EST281945 tomato callus, TAMU Lycopersicon esculentum cDNA clone CLEC36D8, mRNA sequence.
 ACCESSION AM035791
 VERSION AM035791.1 GI:5894547
 KEYWORDS EST.
 SOURCE Lycopersicon esculentum (tomato)
 ORGANISM
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
 REFERENCE 1 (bases 1 to 613)
 Alcala, J., Vrebalov, J., White, R., Matern, A.L., Vision, T., Holt, I.E., Liang, F., Upton, J., Craven, M.B., Bowman, C.L., Ahn, S., Ronning, C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J.
 Generation of ESTs from tomato callus tissue
 Unpublished (1999)
 CONTACT: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.
 Location/Qualifiers
 1..613
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="CLEC36D8"
 /tissue_type="callus"
 /dev_stage="25-40 days old"

```

/lab host="XLI-Blue MRF"
/clone lib="tomato callus, TAMU"
/notes="vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Giovannoni laboratory; cLEC - Cotyledons
of seedlings 7-10 days post-germination were excised, cut
at both ends and placed on MS medium with no selection.
Mixed callus was harvested at 25 and 40 days and included
undifferentiated masses. Tomato Callus EST Library"

```

ORIGIN

```

Query Match 41.7%; Score 245; DB 9; Length 613;
Best Local Similarity 72.5%; Pred. No. 4.4e-58;
Matches 317; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 135 GGTCAAAAGTCACCGAGTATACAGTCGCTGTTTACCTGCTAGTCCCAATATC 194
DB 131 GCTCACCCTTCAACGCTGCAAAATGCTCTTCTTCCCGATTTACAGGCCCAAACTC 190
QY 195 TTCTGATTATATCCAGACTCGAAATTTTCAAGGTGAAGCAATTTGTCAGACCATGGAG 254
DB 191 TCCAGATTTCCTCCTGATGCGAGTTTACAAAGTTGAAGCAATTTCAAGACCTTGGAG 250
QY 255 AATCAGCAAGTTTCACTGCGTTTACTGAAATCGGATTCGAGTGTACTGTTCTGA 314
DB 251 AATTCAACAGGTTTCTTCGGCACTACTCAAAATGGCAATTCGTGCTGCTGTTTCGA 310
QY 315 TGTGAGAGGTTTGTGTCACAGGAGGTTCTACCGAGAGACAGCTGCTCTGAGTTCTC 374
DB 311 TGTGCTGCTGTTTGGCGCCCAAGTGGCTTCACTGAGAGGCAAGCTGCTCTGAATCTC 370
QY 375 GGAAGACAAATTTGTTGTAAGTTAAGTGAATCGTTGTTAAGAAAGACCAAGTGA 434
DB 371 TGAAGACAGCTTTGTTGCAAAAGTTAAATGGAATTTGTTGCAAGAACCCAGTTGA 430
QY 435 ATCTGTAATCAACATTAATGAAGAGAGAGAGAGAGAGAGATTTGTTGATGCGAGAT 494
DB 431 AGGAGTCAATGCCATGATAATGAAGAGGCAAGACTGCTGAAATAGGTGATGGAAGAT 490
QY 495 TTTTGTGTTGCTGTGTCAGATGTCATAGAGTTAGGACAGTGGCTGGGGAGAAAGC 554
DB 491 ATCTTGACTCCCATCTCCGATGTTATAAGAGTTGCGACTGGTGAACCGGAGAAAGC 550
QY 555 AGAGAAGATGACTGGTG 571
DB 551 AGAGAGATGATGGAG 567

```

```

RESULT 14
BI932123
LOCUS
DEFINITION
EST552012 tomato flower, 8 mm to preanthesis buds Lycopersicon
esculentum cDNA clone cTCC22G23 5' end, mRNA sequence.
BI932123
VERSION
KEYWORDS
SOURCE
ORGANISM
Lycopersicon esculentum (tomato)

```

```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 702)

```

```

van der Hoeven, R.S., Bezzerides, J.L., Karamycheva, S.A., Tsai, J.,
Utterback, T., Van Aken, S., Ronning, C.M., Nierman, W., Frazer, C.M.,
Martín, G.B., Giovannoni, J.D. and Tanksley, S.D.

```

```

Generation of ESTs from tomato flower tissue, buds 8 mm -
preanthesis

```

```

Unpublished (2001)
Contact: CUGI

```

```

Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html

```

```

This clone is available through the Clemson University Genomics

```

```

Institute
Seq primer: T3.
Location/Qualifiers
i. 702
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cTCC22G23"
/tissue_type="flower"
/dev_stage="buds 8mm to preanthesis"
/clone_lib="tomato flower, 8 mm to preanthesis buds"
/notes="vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Cornell University; sequencing: The
Institute for Genomic Research; Flower buds and flowers
were taken from greenhouse plants (4-8 wks old, TA496).
They were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

```

ORIGIN

```

Query Match 41.7%; Score 245; DB 12; Length 702;
Best Local Similarity 72.5%; Pred. No. 4.6e-58;
Matches 317; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 135 GGTCAAAAGTCACCGAGTATACAGTCGCTGTTTACCTGCTAGTCCCAATATC 194
DB 172 GCTCACCCTTCAACGCTGTCAAAATGCTCTTCTTCCCGATTTACAGGCCCAAACTC 231
QY 195 TTCTGATTATATTCAGACTCGAAATTTTACAGGTGGAGCAATTTGTCAGACCATGGAG 254
DB 232 TCCAGATTTCGTCCTGATGCGAGTTTTCACAAAGTTGAAGCAATTTCTAAGACCTTGGAG 291
QY 255 AATCAGCAAGTTTTCATCGGCTTTACTGAAATCGGATTCGAGGTGTTACTGTTCTGA 314
DB 232 AATCAGCAAGTTTTCATCGGCTTTACTGAAATCGGATTCGAGGTGTTACTGTTCTGA 351
QY 315 TGTGAGAGGTTTGTGTCACAGGAGGTTCTACCGAGAGACAGCTGGCTGTCGATCTC 374
DB 352 TGTGCTGTTTGGCGCCCAAGTGGCTTGAAGAGGCAAGCTGGCTCTGAATCTC 411
QY 375 GGAAGACAAATTTGTTGCTAAAGTTAAGTGAATCGTTGTTAAGAAAGACCAAGTGA 434
DB 412 TGAAGACACGTTTGTGCAAAAGTTTAAATGGAATTTGTGTCAGCAAGACAGGTTGA 471
QY 435 ATCTGTAATCAACATTAATGAAGAGGCAAGGAGAGAGATTTGTCAGCAAGAT 494
DB 472 AGGAGTCAITGCCATGATTAATGAAGAGGCAAGACTGTTGAAATAGGTGATGGAAGAT 531
QY 495 TTTTGTGTTGCTGTGTCAGATGTCATAGAGTTAGGACAGTGGCTGGGGAGAAAGC 554
DB 532 ATCTTGACTCCCATCTCCGATGTTATAAGAGTTGCGACTGGTGAACCGGAGAAAGC 591
QY 555 AGAGAGATGACTGGTG 571
DB 532 AGAGAGATGATGGAG 608

```

```

RESULT 15
BI932336
LOCUS
DEFINITION

```

```

EST552225 tomato flower, 8 mm to preanthesis buds Lycopersicon
esculentum cDNA clone cTCC22H10 5' end, mRNA sequence.
BI932336
VERSION
KEYWORDS
SOURCE
ORGANISM
Lycopersicon esculentum (tomato)

```

```

Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 711)

```

```

Utterback, T., Van Aken, S., Ronning, C.M., Nierman, W., Frazer, C.M.,

```

```

Generation of ESTs from tomato flower tissue, buds 8 mm -
preanthesis

```

```

Unpublished (2001)
Contact: CUGI

```

```

Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html

```

```

This clone is available through the Clemson University Genomics

```

Martin, G.B., Giovannoni, J.J. and Tanksley, S.D.
Generation of ESTs from tomato flower tissue, buds 8 mm -

preanthesis
Unpublished (2001)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
This clone is available through the Clemson University Genomics
Institute

Seq primer: T3.

Location/Qualifiers

1..71:
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cT0C22H10"
/tissue_type="flower"
/dev_stage="buds 8mm to preanthesis"
/clone_lib="tomato flower, 8 mm to preanthesis buds"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Cornell University; sequencing: The
Institute for Genomic Research; Flower buds and flowers
were taken from greenhouse plants (4-8 wks old, TA496).
They were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

ORIGIN

Query Match	41.7%	Score 245;	DB 12;	Length 711;
Best Local Similarity	72.5%	Pred. No. 4.6e-58;		
Matches 317;	Conservative 0;	Mismatches 120;	Indels 0;	Gaps 0;

QY	135	GGTCACAAAGTCCAGAGTAATAACAGTCGTGTTTACCTGTCGTGTAGTGCCTCAATATC	194
DB	172	GCTCACCTTCAAACGGCTGTCAAAATGCTCTCTTCCGATTATCAGAGCCCAAACTC	231
QY	195	TTCTGATTATATCCAGACTCGAAATTTACAGGTGGAAGCAATTTCTCAGCCATGGAG	254
DB	232	TCCAGATTTTCGCTCCCTGATCGGAAGTTTACAAAGTTGAAGCAATTTCTAAGACCTTGGAG	291
QY	255	AATCCAGCAGGTTTCATCGGCTTTACTGAAAATCGGGATTGAGGTGTTACTGTTCTGA	314
DB	292	AATTCACAGGTTTCTTCGGCACTACTGAAATGGGCAATTCGTGTGTTCTCAGA	351
QY	315	TGTGAGAGGTTTGGTGGCAACAGAGTTCTACCGAGAGACACGGTGGCTCTGAGTTCTC	374
DB	352	TGTTCTGGTGTGTTGGGCCCCAAGTGGCTTGACTGAGAGGCAAGCTGGCTCTGAATTC	411
QY	375	GGAGACAAATTTGTTCTTAAGTTAAGATGGAATCGTTGTTAAGAAAGCAAGTGA	434
DB	412	TGAAGACACGTTGTTGTCAAAAGTTAAATGGAATTTGTCGCAAGACCAAGTTGA	471
QY	435	ATCTGTAATCAACACAAATATTGAAGAGCAAGCAAGAGAGATTGGTGTATGCAAGAT	494
DB	472	AGAGTCATTGCCATGATTAATTGAGAGGCAAGACTGGTGAATAGTGTGGAAGAT	531
QY	495	TTTGTGTTTCCCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG	554
DB	532	ATCTTTGACTCCCATCTCCGATGTTATAGAGTTCCGACTGTTGAACGGGAGATAGGC	591
QY	555	AGAGAGATGACTGGTG	571
DB	592	AGAGAGATGATGGAG	608

Search completed: May 28, 2004, 00:14:54
Job time : 1363.19 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2004, 13:28:59 ; Search time 205.069 Seconds
(without alignments)
12180.972 Million cell updates/sec

Title: US-09-756-541-15
Perfect score: 588
Sequence: 1 ATGGCGCGCTCAATCAACGNA.....GTGATATGCTTTCACCGTCT 588

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 29Jan04: *
1: geneseqn1980s: *
2: geneseqn1990s: *
3: geneseqn2000s: *
4: geneseqn2001as: *
5: geneseqn2001bs: *
6: geneseqn2002s: *
7: geneseqn2003as: *
8: geneseqn2003bs: *
9: geneseqn2003cs: *
10: geneseqn2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	588	100.0	588	4	Aaf58583 Arabidops
2	588	100.0	591	6	Abz14328 Arabidops
3	588	100.0	817	4	Aaf58581 Arabidops
4	285	48.5	594	4	Aaf58584 Ricinus c
5	285	48.5	997	4	Aaf58582 Ricinus c
6	92.8	15.8	96109	4	Aaf28548 Genomic f
7	85.8	14.6	375	8	Ada30024 DNA encod
8	84.6	14.4	110000	2	Continuation (2 of
9	84.4	14.4	110000	2	Continuation (15 o
10	81.4	13.8	339	7	Acf72027 Phototrab
11	81.4	13.8	110000	7	Continuation (53 o
12	81.4	13.8	110000	7	Continuation (4 of
13	79.8	13.6	110000	2	Continuation (4 of
14	77	13.1	363	7	Abz40022 N. gonorr
15	73	12.4	16526	3	Aaa81472 N. mening
16	73	12.4	110000	3	Continuation (7 of
17	73	12.4	172325	3	Aaf21613 Neisseria
18	73	12.4	349980	3	Aaf21612 Neisseria
19	68.8	11.7	336	5	Aah58445 C glutami
20	68.8	11.7	417	5	Aah67222 C glutami
21	68.8	11.7	349980	5	Aah68531 C glutami
22	67.2	11.4	1164	5	Aas88441 DNA encod
23	67.2	11.4	11095	4	Aas46256 DNA encod

ALIGNMENTS

RESULT 1

Aaf58583
ID AAF58583 standard; cDNA; 588 BP.

XX AAF58583;

DT 23-APR-2001 (first entry)

XX Arabidopsis thaliana P-PII cDNA fragment.

XX Arabidopsis thaliana; PII; plant nitrogen regulatory gene; P-PII;
XX nitrogen assimilation; transgenic plant; herbicide screening; ss.

XX Arabidopsis thaliana.

XX US6177275-B1.

XX 23-JAN-2001.

XX 23-JUL-1997; 97US-00899330.

XX 24-JUL-1996; 96US-0022328P.

XX (UNY) UNIV NEW YORK STATE.

XX Coruzzi GM, Lam H, Hsieh M;

XX WPI; 2001-158572/16.

XX P-PSDB; AAB69495.

XX Novel P-PII genes capable of regulating plant nitrogen assimilation,
XX useful for transgenic plant production, and as probes for isolating
XX additional genomic clones having P-PII gene promoters.

XX Claim 2; Col 37-38; 35pp; English.

XX The present sequence encodes a nitrogen regulatory PII protein. Novel
XX plant PII (also called P-PII) nucleotide sequences have been isolated.
XX They are useful for regulating nitrogen assimilation in plants, and in
XX transgenic plant production. They are also used to engineer organisms
XX that overexpress wild-type or mutant P-PII regulatory proteins. P-PII
XX proteins are useful for in vitro screening of herbicides. P-PII
XX nucleotides may be used as probes for isolating additional genomic clones
XX with the promoters of P-PII genes. P-PII promoters are light- and/or
XX sucrose-inducible, and are suitable for genetic engineering of plants

Acf65588 Phototrab
Continuation (47 o
Continuation (2 of
Acf71482 Phototrab
Continuation (16 o
Continuation (2 of
Continuation (3 of
Aaf71296 Corynebact
Continuation (17 o
Abq70750 Listeria
Continuation (16 o
Continuation (33 o
Continuation (33 o
Abn18985 Human ORF
Abz39932 N. gonorr
Aas89065 DNA encod
Aas90406 DNA encod
Aas85755 DNA encod
Aas92604 DNA encod
Ada71938 Rice gene
Aaz99179 T. hydrot
Acf67052 Phototrab

24 64.6 11.0 663 7 ACF65588
25 64.6 11.0 110000 7 ACF67367 46
26 64.6 11.0 110000 7 ACF65388_01
27 61 10.4 339 7 ACF71482
28 60.4 10.3 110000 6 ABQ69245_15
29 60.4 10.3 110000 6 ABQ67195_1
30 60.4 10.3 110000 6 ABQ67195_2
31 59.6 10.1 290 4 AAF71296
32 57.6 9.8 110000 6 ABA90521_16
33 52.4 8.9 2751 6 ABQ70750
34 50.8 8.6 110000 6 ABA03041_15
35 49.6 8.4 110000 4 AAI99682_32
36 49.6 8.4 110000 4 AAI99683_32
37 47 8.0 342 6 ABN18985
38 44.4 7.6 174 7 ABZ39932
39 43.4 7.4 3858 5 AAS89065
40 43.4 7.4 3858 5 AAS90406
41 43.4 7.4 3858 5 AAS85755
42 43.4 7.4 3858 5 AAS92604
43 43 7.3 2000 7 ADA71938
44 42.8 7.3 1970 3 AAZ99179
45 42.2 7.2 674 7 ACF67052

SQ Sequence 588 BP; 164 A; 105 C; 147 G; 172 T; 0 U; 0 Other;
 Query Match 100.0%; Score 588; DB 4; Length 588;
 Best Local Similarity 100.0%; Pred. No. 2.3e-168;
 Matches 588; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGCGCGGTCATGACGAAACCCATCTCAATACTCTCTCGGTTCTATTCTGATCGA 60
 DB 1 ATGCGCGGTCATGACGAAACCCATCTCAATACTCTCTCGGTTCTATTCTGATCGA 60
 QY 61 AAGAACATGCTTCTCTGATGCAATTCGATTGTTCTGATTTCAGACATTCGCGACCA 120
 DB 61 AAGAACATGCTTCTCTGATGCAATTCGATTGTTCTGATTTCAGACATTCGCGACCA 120
 QY 121 TCTTGCCTCGATTGGTCACAAAGTCACCGAGTAAATAACAGTCTGTTTACCTGCTT 180
 DB 121 TCTTGCCTCGATTGGTCACAAAGTCACCGAGTAAATAACAGTCTGTTTACCTGCTT 180
 QY 181 AGTCCCAAAATATCTCTGATTATATTCAGACTGAAATTTTAAAGGTCGAAGCAATT 240
 DB 181 AGTCCCAAAATATCTCTGATTATATTCAGACTGAAATTTTAAAGGTCGAAGCAATT 240
 QY 241 GTGACACATGCGAATCCACAAAGTTCATCGGCTTTTACTGAAATTCGGGATTCGAGT 300
 DB 241 GTGACACATGCGAATCCACAAAGTTCATCGGCTTTTACTGAAATTCGGGATTCGAGT 300
 QY 301 GTTACTGTTCTGATGAGAGGTTTGTGTCACAAAGAGTTCTTACCGAGAGACAGGT 360
 DB 301 GTTACTGTTCTGATGAGAGGTTTGTGTCACAAAGAGTTCTTACCGAGAGACAGGT 360
 QY 361 GGCTCTGAGTCTCGAAGACAAATTTGTTCTTAAAGTTAAGTGAATCGTTTAAAG 420
 DB 361 GGCTCTGAGTCTCGAAGACAAATTTGTTCTTAAAGTTAAGTGAATCGTTTAAAG 420
 QY 421 AAGACCAAGTGGAAATCTGTAATCAACAAATTAATGAAAGCAAGGACAGGAGATT 480
 DB 421 AAGACCAAGTGGAAATCTGTAATCAACAAATTAATGAAAGCAAGGACAGGAGATT 480
 QY 481 GGTGATGGCAAGATTTTGTGTTGCTGTGTCAGATGTCATAAGATTAGGACAGGTGAG 540
 DB 481 GGTGATGGCAAGATTTTGTGTTGCTGTGTCAGATGTCATAAGATTAGGACAGGTGAG 540
 QY 541 CGTGGGAGAAAGCAGAGAGATGCTGTTGATATGCTTTCACCGTCT 588
 DB 541 CGTGGGAGAAAGCAGAGAGATGCTGTTGATATGCTTTCACCGTCT 588

RESULT 2

ID ABZ14328 standard; DNA; 591 BP.
 AC
 XX ABZ14328;

21-JAN-2003 (first entry)

Arabidopsis thaliana stress response regulated gene SEQ ID NO 2133.

Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

Arabidopsis thaliana.

WO200216655-A2.

28-FEB-2002.

24-AUG-2001; 2001WO-US026685.

24-AUG-2000; 2000US-0227866P.

26-JUN-2001; 2001US-0264647P.

22-JUN-2001; 2001US-0300111P.

(SCRI) SCRIPPS RES INST.

(SYGN) SYNGENTA PARTICIPATIONS AG.

XX Harper JP, Kreps J, Wang X, Zhu T;
 XX WPI; 2002-304127/34.
 XX Identifying a stress condition to which a plant cell has been exposed and
 XX producing plants with increased tolerance to these abiotic stresses.
 PT Claim 144; SEQ ID NO 2133; 577bp + Sequence Listing; English.

XX The invention relates to identifying a stress condition to which a plant
 CC cell has been exposed, comprising: (a) contacting nucleic acid
 CC representative of expressed polynucleotides in the plant cell with an
 CC array or probes representative of the plant cell genome; and (b)
 CC detecting a profile of expressed polynucleotides in the plant cell
 CC characteristic of a stress response. The method is useful in the
 CC production of transgenic plants, cells and seeds and in producing plants
 CC with increased tolerance to abiotic stress. The present sequence is that
 CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
 CC in methods of the invention. Note: The sequence data for this patent is
 CC not represented in the printed specification but is based on sequence
 CC information supplied to Derwent by the European Patent Office

XX Sequence 591 BP; 165 A; 105 C; 148 G; 173 T; 0 U; 0 Other;

Query Match 100.0%; Score 588; DB 6; Length 591;

Best Local Similarity 100.0%; Pred. No. 2.3e-168;
 Matches 588; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCGCGGTCATGACGAAACCCATCTCAATACTCTCTCGGTTCTATTCTGATCGA 60
 DB 1 ATGCGCGGTCATGACGAAACCCATCTCAATACTCTCTCGGTTCTATTCTGATCGA 60
 QY 61 AAGAACATGCTTCTCTGATGCAATTCGATTGTTCTGATTTCAGACATTCGCGACCA 120
 DB 61 AAGAACATGCTTCTCTGATGCAATTCGATTGTTCTGATTTCAGACATTCGCGACCA 120
 QY 121 TCTTGCCTCGATTGGTCACAAAGTTCATCGGCTTTTACTGAAATTCGGGATTCGAGT 180
 DB 121 TCTTGCCTCGATTGGTCACAAAGTTCATCGGCTTTTACTGAAATTCGGGATTCGAGT 180
 QY 181 AGTCCCAAAATATCTCTGATTATATTCAGACTGAAATTTTAAAGGTCGAAGCAATT 240
 DB 181 AGTCCCAAAATATCTCTGATTATATTCAGACTGAAATTTTAAAGGTCGAAGCAATT 240
 QY 241 GTGACACATGCGAATCCACAAAGTTCATCGGCTTTTACTGAAATTCGGGATTCGAGT 300
 DB 241 GTGACACATGCGAATCCACAAAGTTCATCGGCTTTTACTGAAATTCGGGATTCGAGT 300
 QY 301 GTTACTGTTCTGATGAGAGGTTTGTGTCACAAAGAGTTCTTACCGAGAGACAGGT 360
 DB 301 GTTACTGTTCTGATGAGAGGTTTGTGTCACAAAGAGTTCTTACCGAGAGACAGGT 360
 QY 361 GGCTCTGAGTCTCGAAGACAAATTTGTTCTTAAAGTTAAGTGAATCGTTTAAAG 420
 DB 361 GGCTCTGAGTCTCGAAGACAAATTTGTTCTTAAAGTTAAGTGAATCGTTTAAAG 420
 QY 421 AAGACCAAGTGGAAATCTGTAATCAACAAATTAATGAAAGCAAGGACAGGAGATT 480
 DB 421 AAGACCAAGTGGAAATCTGTAATCAACAAATTAATGAAAGCAAGGACAGGAGATT 480
 QY 481 GGTGATGGCAAGATTTTGTGTTGCTGTGTCAGATGTCATAAGATTAGGACAGGTGAG 540
 DB 481 GGTGATGGCAAGATTTTGTGTTGCTGTGTCAGATGTCATAAGATTAGGACAGGTGAG 540
 QY 541 CGTGGGAGAAAGCAGAGAGATGCTGTTGATATGCTTTCACCGTCT 588
 DB 541 CGTGGGAGAAAGCAGAGAGATGCTGTTGATATGCTTTCACCGTCT 588

RESULT 3

AAF58581
 ID AAF58581 standard; cDNA; 817 BP.

XX AC AAF58581;
 XX DT 23-APR-2001 (first entry)
 XX DE Arabidopsis thaliana P-PII cDNA.
 XX KW Arabidopsis thaliana; PII; plant nitrogen regulatory gene; P-PII;
 XX KW nitrogen assimilation; transgenic plant; herbicide screening; ss.
 XX OS Arabidopsis thaliana.
 XX PN US6177275-B1.
 XX PD 23-JAN-2001.
 XX PF 23-JUL-1997; 97US-00899330.
 XX PR 24-JUL-1996; 96US-0022328P.
 XX FA (UJNY) UNIV NEW YORK STATE.
 XX PI Coruzzi GM, Lam H, Hsieh M;
 XX DR WPI; 2001-158572/16.
 XX DR P-PSDB; AAB69495.
 XX Novel P-PII genes capable of regulating plant nitrogen assimilation,
 PT useful for transgenic plant production, and as probes for isolating
 PT additional genomic clones having P-PII gene promoters.
 XX Claim 2; Fig 12; 35pp; English.
 XX The present sequence encodes a nitrogen regulatory PII protein. Novel
 CC plant PII (also called P-PII) nucleotide sequences have been isolated.
 CC They are useful for regulating nitrogen assimilation in plants, and in
 CC transgenic plant production. They are also used to engineer organisms
 CC that overexpress wild-type or mutant P-PII regulatory proteins. P-PII
 CC proteins are useful for in vitro screening of herbicides. P-PII
 CC nucleotides may be used as probes for isolating additional genomic clones
 CC with the promoters of P-PII genes. P-PII promoters are light- and/or
 CC sucrose-inducible, and are suitable for genetic engineering of plants
 XX Sequence 817 BP; 242 A; 137 C; 188 G; 250 T; 0 U; 0 Other;
 SQ Query Match 100.0%; Score 588; DB 4; Length 817;
 Best Local Similarity 100.0%; Pred. No. 2.6e-168;
 Matches 588; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGCGGGTCAATGACGAAACCCATCTCAATACTTCTCTGGTTCTTATCTGATCGA 60
 DB 33 ATGGCGGGTCAATGACGAAACCCATCTCAATACTTCTCTGGTTCTTATCTGATCGA 92
 QY 61 AAGAACATTGCTTCTCTGATTGCAATTCGATTCTCTCGATTCCGACCA 120
 DB 93 AAGAACATTGCTTCTCTGATTGCAATTCGATTCTCTCGATTCCGACCA 152
 QY 121 TCTTGCTCGATTGGTCAACAAGTCAACGAGTAATAACAGTGGTTTACCTGCGTT 180
 DB 153 TCTTGCTCGATTGGTCAACAAGTCAACGAGTAATAACAGTGGTTTACCTGCGTT 212
 QY 181 AGTCCGCAATATCTTCTGATTATATTCAGATCCAGATCGAAATTTACAGGTGGAACATT 240
 DB 213 AGTCCGCAATATCTTCTGATTATATTCAGATCCAGATCGAAATTTACAGGTGGAACATT 272
 QY 241 GTCAGACCATCGAGATCCAGCAAGTTTCATCGGCTTTACTGAAATCGGATTCGAGT 300
 DB 273 GTCAGACCATCGAGATCCAGCAAGTTTCATCGGCTTTACTGAAATCGGATTCGAGT 332
 QY 301 GTTACTGTTTCTGATGTGAGAGGGTTTGGTGCAACAGGAGGTTTACCGAGAGACACCGT 360
 DB 333 GTTACTGTTTCTGATGTGAGAGGGTTTGGTGCAACAGGAGGTTTACCGAGAGACACCGT 392

QY 361 GGCTCTGAGTTCTCGGACAGCAAAATTTGTTGCTTAAGTTAAGATGGAATCGTTGTTAAG 420
 DB 393 GGCTCTGAGTTCTCGGACAGCAAAATTTGTTGCTTAAGTTAAGATGGAATCGTTGTTAAG 452
 QY 421 AAGACCAAGTGAATCTGTAATCAACACAATAATTGAAGGAGCAAGGACAGGAGATT 480
 DB 453 AAGACCAAGTGAATCTGTAATCAACACAATAATTGAAGGAGCAAGGACAGGAGATT 512
 QY 481 GGTGATGGCAAGATTTTGTGCTGCTGTGTGATGTCATTAAGATTTAGGACAGGTGAG 540
 DB 513 GGTGATGGCAAGATTTTGTGCTGCTGTGTGATGTCATTAAGATTTAGGACAGGTGAG 572
 QY 541 CGTGGGGGAGGACAGGAGATGACTGCTGATGATGCTTTCACCGTCT 588
 DB 573 CGTGGGGGAGGACAGGAGATGACTGCTGATGATGCTTTCACCGTCT 620

RESULT 4
 AAF58584
 ID AAF58584 standard; cDNA; 594 BP.
 XX AC AAF58584;
 XX DT 23-APR-2001 (first entry)
 XX DE Ricinus communis P-PII cDNA fragment.
 XX KW Castor bean; PII; plant nitrogen regulatory gene; P-PII;
 XX KW nitrogen assimilation; transgenic plant; herbicide screening; ss.
 XX OS Ricinus communis.
 XX PN US6177275-B1.
 XX PD 23-JAN-2001.
 XX PF 23-JUL-1997; 97US-00899330.
 XX PR 24-JUL-1996; 96US-0022328P.
 XX FA (UJNY) UNIV NEW YORK STATE.
 XX PI Coruzzi GM, Lam H, Hsieh M;
 XX DR WPI; 2001-158572/16.
 XX DR P-PSDB; AAB69495.
 XX Novel P-PII genes capable of regulating plant nitrogen assimilation,
 PT useful for transgenic plant production, and as probes for isolating
 PT additional genomic clones having P-PII gene promoters.
 XX Claim 2; Col 37-38; 35pp; English.
 XX The present sequence encodes a nitrogen regulatory PII protein. Novel
 CC plant PII (also called P-PII) nucleotide sequences have been isolated.
 CC They are useful for regulating nitrogen assimilation in plants, and in
 CC transgenic plant production. They are also used to engineer organisms
 CC that overexpress wild-type or mutant P-PII regulatory proteins. P-PII
 CC proteins are useful for in vitro screening of herbicides. P-PII
 CC nucleotides may be used as probes for isolating additional genomic clones
 CC with the promoters of P-PII genes. P-PII promoters are light- and/or
 CC sucrose-inducible, and are suitable for genetic engineering of plants
 XX Sequence 594 BP; 149 A; 117 C; 157 G; 171 T; 0 U; 0 Other;
 SQ Query Match 48.5%; Score 285; DB 4; Length 594;
 Best Local Similarity 70.3%; Pred. No. 3.4e-76;
 Matches 397; Conservative 0; Mismatches 165; Indels 3; Gaps 1;
 QY 7 GCGTCAATGACGAAACCCATCTCAATACTTCTCTCGTTTCTTCTGATCGAAAGAAC 66
 DB 7 GAGGCTACTGGAACCTGGGCTTGCTCACTCTCTCTTCTTCTTAATAACATCAAGAGAA 66

XX 28-DEC-2000.
 XX 16-JUN-2000; 2000WO-US016649.
 XX 18-JUN-1999; 99US-0140121P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Lagace RE, Patterson C, Berg XL;
 XX WPI; 2001-041427/05.
 XX Genomic library for identifying diagnostic and therapeutic compositions,
 XX and for identifying virulence factors, regulatory elements and drug
 XX targets, comprises Moraxella catarrhalis nucleic acids.
 XX Claim 1; Page 345-368; 545pp; English.
 XX The present invention relates to a Moraxella catarrhalis genomic library
 XX comprising of a combination of 41 nucleic acid molecules (see AAF28514-
 XX AAF28554). The library has a number of uses described in the
 XX specification e.g. is useful for identifying diagnostic and therapeutic
 XX compositions. M. catarrhalis (Moraxella catarrhalis) is a large
 XX aerobic, gram-negative diplococcus, normally found among the bacterial
 XX flora of human upper airways. M. catarrhalis is known to cause acute,
 XX localised infections such as otitis media, sinusitis and bronchopulmonary
 XX infection and life-threatening, systemic diseases including endocarditis
 XX and meningitis
 XX Sequence 96109 BP; 28783 A; 18910 C; 20341 G; 28075 T; 0 U; 0 Other;
 Query Match 15.8%; Score 92.8; DB 4; Length 96109;
 Best Local Similarity 57.8%; Pred. No. 8.8e-17;
 Matches 185; Conservative 0; Mismatches 132; Indels 3; Gaps 1;
 QY 235 GCAATTGTCAGACCATGAGATCCAGCAAGTTTCATCGGCTTTACTGAAATCGGATT 294
 DB 12953 GCAATTATCAAGCCGTTTAAACTCGATGATGCGAGAAAGCACTCTCAGAAATTCGCGTC 13012
 QY 295 CGAGGTGTTACTGTTCTGATGTGAGAGGTTTGGTGCACAGGAGGTTCTACCGAGAGA 354
 DB 13013 AATGATATCCGTCACGTAAGTAAAGCTTGTGCGCAAAAGGTCATACCGAGATG 13072
 QY 355 CACGGTGGCTCTGAGTTCTCGAGACACAAATTTGTTGCTAAAGTTAAGATGGAATCGTT 414
 DB 13073 TATCGTGGGGGGAATATGTTGTTGAT---TTTACCACAAATTAATAATTCAGATAGCA 13129
 QY 415 GTTAAGAAGACCAAGTCGATCTGTATATCAACACAAATTAATGAAGGACAGGACNGA 474
 DB 13130 TGTGATGATGATGGTTGATTCATATTTAGTCAATTCATTAAGTTTCAAAATACAGT 13189
 QY 475 GAGATTGCTGATGCAAGATTTTCTGTTTGTGCTGTCAGATGTCATAGAGTTAGGACA 534
 DB 13190 AAATGCTGATGGTAAGATTTTGTAGTCCGCTGAGCGTGTATTCGCAATTCGAACT 13249
 QY 535 GGTGAGCGTGGGAGAAAGC 554
 DB 13250 GCGCAATTTGATGAAAGTGC 13269
 RESULT 7
 ADA30024
 ID ADA30024 standard; DNA; 375 BP.
 XX
 XX ADA30024;
 XX
 XX 20-NOV-2003 (first entry)
 XX DNA encoding Acinetobacter baumannii protein #1311.
 XX ds; gene; Acinetobacter baumannii; bacterial disease; antibacterial;
 XX vaccine; plant biocontrol agent.

XX Acinetobacter baumannii.
 XX US562958-B1.
 XX 13-MAY-2003.
 XX 04-JUN-1999; 99US-00328352.
 XX 09-JUN-1998; 98US-0089701P.
 XX (GENO-) GENOME THERAPEUTICS CORP.
 XX Breton G, Bush D;
 XX WPI; 2003-576092/54.
 XX P-PSDB; ADA34150.
 XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 XX for diagnosing a bacterial disease, as components of antibacterial
 XX vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 XX plants.
 XX Example; SEQ ID NO 1311; 328pp; English.
 XX The invention relates to isolated Acinetobacter baumannii nucleic acids.
 XX The A. baumannii nucleic acids and polypeptides are useful as reagents
 XX for diagnosing a bacterial disease, as components of antibacterial
 XX vaccines, as targets for antibacterial drugs, to detect the presence of
 XX A. baumannii and other Acinetobacter species in a sample, in screening
 XX compounds for the ability to interfere with the A. baumannii life cycle
 XX or to inhibit A. baumannii infection, and as biocontrol agents for
 XX plants. The present sequence represents DNA encoding an A. baumannii
 XX protein.
 XX Sequence 375 BP; 122 A; 64 C; 93 G; 96 T; 0 U; 0 Other;
 Query Match 14.6%; Score 85.8; DB 8; Length 375;
 Best Local Similarity 57.4%; Pred. No. 1.1e-15;
 Matches 175; Conservative 0; Mismatches 127; Indels 3; Gaps 1;
 QY 235 GCAATTGTCAGACCATGAGAAATCCAGCAAGTTTCATCGGCTTTACTGAAATCGGATT 294
 DB 52 GCAATTGTAACCCGTTTAAATGATGATGCGTGAAGCACTCTCTGACATTGGTA 111
 QY 295 CGAGGTGTTACTGTTCTGATGTGAGAGGTTTGGTGCACAGGAGGTTCTACCGAGAGA 354
 DB 112 CAAGGATTAACCGTAACCTGAAGTTAAAGGTTTGGTCTGTCACAAAGSACATACAGAACTT 171
 QY 355 CACGGTGGCTCTGAGTTCTCGGAAGACAAATTTGTTGCTTAAAGTTAAGATGGAATCGTT 414
 DB 172 TACCGCGCGCTGAGTATGTTGTTGAT---TTCCTTACCTAAAGTAAATTCGAATTCG 228
 QY 415 GTTAAGAAGACCAAGTGGAAATCTGTAATCAACACAAATTAATGAAGGACAGGACAGGA 474
 DB 229 ATTAGTGAATGATGTCGCAAGGTTAATGAGTCAATACACGTCGTGGCAAGCACTGGA 288
 QY 475 GAGATTGCTGATGCAAGATTTTGTGTTTGTGCTGTCAGATGTCATAGAGTTAGGACA 534
 DB 289 AAAATCGCGAGCGGTAAGATTTTGTGACTAATCTCGAACAAGTCATCCGTATCCGTACA 348
 QY 535 GGTGA 539
 DB 349 GGTGA 353
 RESULT 8
 AAV21209_01
 Continuation (2 of 17) of AAV21209 from base 100001 (Methanococcus jannaschii circular c
 WP Sequence split into 17 fragments LOCUS AAV21209 Accession Aav21209
 WP Fragment Name Begin End
 WP AAV21209_00 1 110000
 WP AAV21209_01 100001 210000

WP	AAV21209_02	200001	310000	
WP	AAV21209_03	300001	410000	
WP	AAV21209_04	400001	510000	
WP	AAV21209_05	500001	610000	
WP	AAV21209_06	600001	710000	
WP	AAV21209_07	700001	810000	
WP	AAV21209_08	800001	910000	
WP	AAV21209_09	900001	1010000	
WP	AAV21209_10	1000001	1110000	
WP	AAV21209_11	1100001	1210000	
WP	AAV21209_12	1200001	1310000	
WP	AAV21209_13	1300001	1410000	
WP	AAV21209_14	1400001	1510000	
WP	AAV21209_15	1500001	1610000	
WP	AAV21209_16	1600001	1664976	

Query Match 14.4%; Score 84.6; DB 2; Length 110000;
Best Local Similarity 56.0%; Pred. No. 2.9e-14;
Matches 181; Conservative 0; Mismatches 139; Indels 3; Gaps 1;

QY	217	AAATTTTACAGGTGGAGCAATGTCAGACCATCGAGATCCAGCAAGTTTCATCGGCT	276
Db	80891	AAATGGAANAAGTGTGAAGCAATCATAGACCCGGAGAGGTGGAGATTGTTAAAGAGGCT	80950
QY	277	TTACTGAAATCCGATTCGAGGTGTACTGTGTTCTGATGTGAGAGGGTTGGTGACAA	336
Db	80951	TTGTCGATCTGATATGTTGGAATGACGTGTAGTCAGGTTAAGGGTAGGGAGTTCAA	81010
QY	337	GGAGGTTCACCGAGAGACACGGTGGCTCTGAGTCTCCGGAAGACAAATTTGGTCTAAA	396
Db	81011	GGTGGAAATAGTTCGAGAGGTATAGGGGAGAGAGTATATTGTTGAT---TTAATCCAAAG	81067
QY	397	GTTAAGATGAAATCGTTGTTTAAAGAAACCAAGTCGAAATCTGTAATCAACAATAAT	456
Db	81068	GTTAAGATGAAATCGTTGTTTAAAGAGAGAGAGTGTGATATGTTATGATATATATGC	81127
QY	457	GAAGGACAGGACAGAGAGATTGGTGATGGCAAGATTTTGTGTCCTGTGTCAGAT	516
Db	81128	GAGAAATGCAAGAACAGCAAAACCCAGGAGATGGAATAATCTTCGTATACCATAGAAAGA	81187
QY	517	GTCNTAAGAGTTAGGACAGGTGA	539
Db	81188	GTCGTAAGAGTTAAGACAAAAGA	81210

RESULT 9
AAV21209 14/C
Continuation (15 of 17) of AAV21209 from base 1400001 (Methanococcus jannaschii circular genome)
WP Sequence split into 17 fragments, LOCUS AAV21209 Accession AAV21209

WP	Fragment Name	Begin	End
WP	AAV21209_00	1	110000
WP	AAV21209_01	100001	210000
WP	AAV21209_02	200001	310000
WP	AAV21209_03	300001	410000
WP	AAV21209_04	400001	510000
WP	AAV21209_05	500001	610000
WP	AAV21209_06	600001	710000
WP	AAV21209_07	700001	810000
WP	AAV21209_08	800001	910000
WP	AAV21209_09	900001	1010000
WP	AAV21209_10	1000001	1110000
WP	AAV21209_11	1100001	1210000
WP	AAV21209_12	1200001	1310000
WP	AAV21209_13	1300001	1410000
WP	AAV21209_14	1400001	1510000
WP	AAV21209_15	1500001	1610000
WP	AAV21209_16	1600001	1664976

Query Match 14.4%; Score 84.4; DB 2; Length 110000;
Best Local Similarity 54.2%; Pred. No. 3.4e-14;
Matches 194; Conservative 0; Mismatches 161; Indels 3; Gaps 1;

198 TGATTATATCCAGACTCGAAATTTTACAAGGTGGAGCAATTCGTCAGACCATGGAGAT 257

CC luminescens. Cells transformed to express the genes are useful for
CC recombinant production of the proteins, particularly toxins and
CC antibiotics useful as insecticides, bactericides and fungicides. The
CC genes, proteins, vectors containing the genes and Ab are also useful
CC therapeutically to treat microbial infection by bacteria or fungi that
CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
CC biopesticides. Other uses of the genes and the proteins are as virulence
CC factors and for identifying targets of human diseases for which P.
CC luminescens is a model (particularly plague and whooping cough). This
CC sequence represents one of the isolated P. luminescens genes
XX
SQ Sequence 339 BP; 105 A; 51 C; 96 G; 87 T; 0 U; 0 Other;

Query Match 13.8%; Score 81.4; DB 7; Length 339;
Best Local Similarity 55.0%; Pred. No. 2.3e-14;
Matches 182; Conservative 0; Mismatches 146; Indels 3; Gaps 1;

QY 224 ACAAGTGGAGCAATTCGACCATGAGATCCAGCAAGTTTCATCGGCTTTACTGA 283
DB 5 AAAAGATTGATCGGATATCAACCTTTCAATTAGATGATGCGTGAAGCTCTGGCGG 64
QY 284 AAATCGGATTCGAGGTGTTACTGTTCTGATGTGAGAGGGTTTGGTGACAAAGAGGTT 343
DB 65 AAGTGGTATCACCGGAATGACAGTAACAGAGGTGAAGGTTTGGGCGCCAAAAGGTC 124
QY 344 CTACCGAGAGACACGGTGGCTCTGAGTTCGGAAGACAAATTTGTTGCTAAAGTTAAGA 403
DB 125 ATACAGAGCTGTATCGCGGTGAGAAATATATGTTGGAT---TTTCTGCCAAAAGTGAAA 181
QY 404 TGGAAATCGTTGTTAAGAAAGACCAAGTGAATCTGTAATCAACAAATTAAGAGGAG 463
DB 182 TAGAAATTGTCGTCGAGATGATATTTGTCGATACCTGTGTTGAACCATATGCGACGG 241
QY 464 CAAGCAGAGGAGATGTTGTCGATGCGACAAATTTTGTGCTGTGTCAGATGTCATAA 523
DB 242 CACAGACCGGGAATTCGTGATGTTGAATATTTGTTATGATGTAGACGTTGTGTC 301
QY 524 GAGTTAGGACAGGTGAGCGTGGGAGAAAGC 554
DB 302 GTATCCGACCGGTGAGCAGGATGAAGAGGC 332

RESULT 11
ACF67367_52
Continuation (53 of 57) of ACF67367 from base 5200001 (Photobacterium luminescens nucleotide
WP Sequence split into 57 fragments LOCUS ACF67367 Accession ACF67367

WP	Fragment Name	Begin	End
WP	ACF67367_53	1	110000
WP	ACF67367_54	100001	210000
WP	ACF67367_55	200001	310000
WP	ACF67367_56	300001	410000
WP	ACF67367_57	400001	510000
WP	ACF67367_58	500001	610000
WP	ACF67367_59	600001	710000
WP	ACF67367_60	700001	810000
WP	ACF67367_61	800001	910000
WP	ACF67367_62	900001	1010000
WP	ACF67367_63	1000001	1110000
WP	ACF67367_64	1100001	1210000
WP	ACF67367_65	1200001	1310000
WP	ACF67367_66	1300001	1410000
WP	ACF67367_67	1400001	1510000
WP	ACF67367_68	1500001	1610000
WP	ACF67367_69	1600001	1710000
WP	ACF67367_70	1700001	1810000
WP	ACF67367_71	1800001	1910000
WP	ACF67367_72	1900001	2010000
WP	ACF67367_73	2000001	2110000
WP	ACF67367_74	2100001	2210000
WP	ACF67367_75	2200001	2310000
WP	ACF67367_76	2300001	2410000
WP	ACF67367_77	2400001	2510000
WP	ACF67367_78	2500001	2610000

WP	Fragment Name	Begin	End
WP	ACF67367_26	2600001	2710000
WP	ACF67367_27	2700001	2810000
WP	ACF67367_28	2800001	2910000
WP	ACF67367_29	2900001	3010000
WP	ACF67367_30	3000001	3110000
WP	ACF67367_31	3100001	3210000
WP	ACF67367_32	3200001	3310000
WP	ACF67367_33	3300001	3410000
WP	ACF67367_34	3400001	3510000
WP	ACF67367_35	3500001	3610000
WP	ACF67367_36	3600001	3710000
WP	ACF67367_37	3700001	3810000
WP	ACF67367_38	3800001	3910000
WP	ACF67367_39	3900001	4010000
WP	ACF67367_40	4000001	4110000
WP	ACF67367_41	4100001	4210000
WP	ACF67367_42	4200001	4310000
WP	ACF67367_43	4300001	4410000
WP	ACF67367_44	4400001	4510000
WP	ACF67367_45	4500001	4610000
WP	ACF67367_46	4600001	4710000
WP	ACF67367_47	4700001	4810000
WP	ACF67367_48	4800001	4910000
WP	ACF67367_49	4900001	5010000
WP	ACF67367_50	5000001	5110000
WP	ACF67367_51	5100001	5210000
WP	ACF67367_52	5200001	5310000
WP	ACF67367_53	5300001	5410000
WP	ACF67367_54	5400001	5510000
WP	ACF67367_55	5500001	5610000
WP	ACF67367_56	5600001	5648894

Query Match 13.8%; Score 81.4; DB 7; Length 110000;
Best Local Similarity 55.0%; Pred. No. 2.8e-13;
Matches 182; Conservative 0; Mismatches 146; Indels 3; Gaps 1;

QY 224 ACAGGTGGAGCAATTCGACCATGAGATCCAGCAAGTTTCATCGGCTTTACTGA 283
DB 49307 AAAAGATTGATCGGATATCAACCTTTCAATTAGATGATGCGTGAAGCTCTGGCGG 49366
QY 284 AAATCGGATTCGAGGTGTTACTGTTCTGATGTGAGAGGGTTTGGTGCAAAAGAGGTT 343
DB 49367 AAGTGGTATCACCGGAATGACAGTAACAGAGGTGAAGGTTTGGGCGCCAAAAGGTC 49426
QY 344 CTACCGAGAGACACGGTGGCTCTGAGTTCGGAAGACAAATTTGTTGCTAAAGTTAAGA 403
DB 49427 ATACAGAGCTGTATCGCGGTGAGAAATATATGTTGGAT---TTTCTGCCAAAAGTGAAA 49483
QY 404 TGGAAATCGTTGTTAAGAAAGACCAAGTGAATCTGTAATCAACAAATTAAGAGGAG 463
DB 49484 TAGAAATTGTCGTCGAGATGATATTTGTCGATACCTGTGTTGAACCATATGCGACGG 49543
QY 464 CAAGCAGAGGAGATGTTGTCGATGCGACAAATTTTGTGCTGTGTCAGATGTCATAA 523
DB 49544 CACAGACCGGGAATTCGTGATGTTGAATATTTGTTATGATGTAGACGCTGTGTC 49603
QY 524 GAGTTAGGACAGGTGAGCGTGGGAGAAAGC 554
DB 49604 GTATCCGACCGGTGAGCAGGATGAAGAGGC 49634

RESULT 12
ACF65387_3
Continuation (4 of 7) of ACF65387 from base 300001 (Photobacterium luminescens nucleotide
WP Sequence split into 7 fragments LOCUS ACF65387 Accession ACF65387

WP	Fragment Name	Begin	End
WP	ACF65387_0	1	110000
WP	ACF65387_1	100001	210000
WP	ACF65387_2	200001	310000
WP	ACF65387_3	300001	410000
WP	ACF65387_4	400001	510000
WP	ACF65387_5	500001	610000
WP	ACF65387_6	600001	696798

```
Query Match      13.8%; Score 81.4; DB 7; Length 110000;
Best Local Similarity 55.0%; Pred. No. 2.8e-13;
Matches 182; Conservative 0; Mismatches 146; Indels 3; Gaps 1;

QY 224 ACAGGTGGAGCAATGTGTCAGACCATGGAGATCCAGCAAGTTTCATCGGCTTACTGA 283
DB 88821 AAAAGATTGATCGATTATCAAACTTCAATAGATGATGCGTGAGAGCTCTCGCGG 88880

QY 284 AAATCGGGATTGAGGTGTTACTGTTCTGATGTGAGAGGTTTGGTGCAAGAGGTT 343
DB 88881 AAGTGGGTATCCCGGAGTACAGATACAGAGTGGAAGGTTTGGGCGCCAAAAGGTC 88940

QY 344 CTACCCGAGACACGGTGGCTCTGAGTCTCGGAAGACAAATTTGTTGCTTAAAGTTAAGA 403
DB 88941 ATACAGAGCTGTATCGCGGTGCAATATATGCTGGAT---TTCTGCCAAAAGTGAAGA 88997

QY 404 TGGAAATCGTTGTTAAGAAAGACCAAGTGAATCTGTAATCAACAAATAATTGAAGGAG 463
DB 88998 TAGAAATTCGTGCGCCAGATGATTTGTCATACCTGTTGTAACCAATATGCAAGCGG 89057

QY 464 CAAAGCAGACAGAGATGTTGTCAGGCAAGATTTTGTGTTTCCCTGTCAGATGTCATAA 523
DB 89058 CACAGCCGGGAAATCGGTGATGTTAAATATTTGTAATTTGATGTAGCACGTTGTCG 89117

QY 524 GAGTTAGGACAGTGTGAGCGTGGGAGAAAGC 554
DB 89118 GTATCCGCACCGGTGAGCAGGATGAAGGCG 89148

RESULT 13
AAT42063_03
Continuation (4 of 19) of AAT42063 from base 320001 (Haemophilus influenzae complete gen
WP Sequence split into 19 fragments LOCUS AAT42063 Accession Aat42063
W2 Fragment Name      Begin      End
WP AAT42063_00        1      110000
WP AAT42063_01       100001    210000
WP AAT42063_02       200001    310000
WP AAT42063_03       300001    410000
WP AAT42063_04       400001    510000
WP AAT42063_05       500001    610000
WP AAT42063_06       600001    710000
WP AAT42063_07       700001    810000
WP AAT42063_08       800001    910000
WP AAT42063_09       900001   1010000
WP AAT42063_10      1000001   1110000
WP AAT42063_11      1100001   1210000
WP AAT42063_12      1200001   1310000
WP AAT42063_13      1300001   1410000
WP AAT42063_14      1400001   1510000
WP AAT42063_15      1500001   1610000
WP AAT42063_16      1600001   1710000
WP AAT42063_17      1700001   1810000
WP AAT42063_18      1800001   1830121

Query Match      13.8%; Score 79.8; DB 2; Length 110000;
Best Local Similarity 54.7%; Pred. No. 8.4e-13;
Matches 181; Conservative 0; Mismatches 147; Indels 3; Gaps 1;

QY 224 ACAGGTGGAGCAATGTGTCAGACCATGGAGATCCAGCAAGTTTCATCGGCTTACTGA 283
DB 63919 AAAAATCGAGCAATGATTAACCTTTAAATTAGACGATGTCGAGAGAGTCTTTCAG 63978

QY 284 AAATCGGGATTGAGGTGTTACTGTTCTGATGTGAGAGGTTTGGTGCAAGAGGTT 343
DB 63979 AATTTGCTATTTTCAGGTATGACATCAAGATGACGCGGATTTGGTTCGTCAAAAGGTC 64038

QY 344 CTACCCGAGACACGGTGGCTCTGAGTCTCGGAAGACAAATTTGTTGCTTAAAGTTAAGA 403
DB 64039 ATACAGAACTTTATCGTGGTGGCAATATATGCTGGAT---TTTCTGCCAAAAGTGAAT 64095

QY 404 TGGAAATCGTTGTTAAGAAAGACCAAGTGAATCTGTAATCAACAAATAATTGAAGGAG 463
```

```
DB 64096 TGGAGTGGTAGTTCTCTGATGAGCTTGTGATCAATGTAATTGAAGCGATTATTGAAACGG 64155
QY 464 CAAAGCAGACAGAGATGTTGTCAGTGGCAAGATTTTGTGTTTGCCTGTCAGATGTCATAA 523
DB 64156 CACAAACAGGTAAATCGGTGACGCGCAAAATTTTGTGTTTATCATCGTGTGAGAGCCATCC 64215

QY 524 GAGTTAGGACAGTGTGAGCGTGGGAGAAAGC 554
DB 64216 GCATTGCGACGGCGAGAAAACGAGATGC 64246

RESULT 14
ABZ40022
ID ABZ40022 standard; DNA; 363 BP.
XX AC
XX ABZ40022;
XX 07-MAR-2003 (first entry)
XX N. gonorrhoeae nucleotide sequence SEQ ID 4633.
XX Antibacterial; infection; vaccine; gene therapy; gene; ds.
XX Neisseria gonorrhoeae.
XX WO200279243-A2.
XX PN
XX PD
XX PF
XX 12-FEB-2002; 2002WO-IB002069.
XX PR
XX 12-FEB-2001; 2001GB-00003424.
XX PA (CHIR-) CHIRON SPA.
XX PI Fontana MR, Pizza M, Massignani V, Monaci E;
XX WPI; 2003-058415/05.
XX P-PSDB; ABP79052.
XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a
XX medicament for treating or preventing N. gonorrhoeae infection.
XX Disclosure; Page 528; 815pp; English.
XX The present invention relates to proteins from Neisseria gonorrhoeae.
XX Also disclosed are the nucleic acid molecules encoding the proteins and
XX antibodies that specifically bind to the proteins. The composition
XX comprising the protein, nucleic acid or antibody is useful for the
XX manufacture of a medicament for treating or preventing N. gonorrhoeae
XX infection, this may be in the form of a vaccine or gene therapy.
XX Sequences given in records ABZ37706-ABZ42016 represent nucleic acid
XX molecules of the invention
XX SQ Sequence 363 BP; 84 A; 77 C; 123 G; 79 T; 0 U; 0 Other;

Query Match      13.1%; Score 77; DB 7; Length 363;
Best Local Similarity 54.1%; Pred. No. 5.3e-13;
Matches 180; Conservative 0; Mismatches 150; Indels 3; Gaps 1;

QY 224 ACAGGTGGAGCAATGTGTCAGACCATGGAGATCCAGCAAGTTTCATCGGCTTACTGA 283
DB 32 AAAAATCGAGCGGATTTGTCAAACCGTTCAAGCTCGACGCTGCGGAGCGGCTTGACGG 91

QY 284 AAATCGGGATTGAGGTGTTACTGTTCTGATGTGAGAGGTTTGGTGCAAGAGGTT 343
DB 92 AAATCGCGATTATCGCGGATGACCGTTCAGCGAGGTCAAAAGGTTTCGGGAGGCAAGAGGCG 151

QY 344 CTACCCGAGACACGGTGGCTCTGAGTCTCGGAAGACAAATTTGTTGCTTAAAGTTAAGA 403
DB 152 ATACGAAATCTATCGCGGTGCGGAATACGCGCGTCTGAT---TTCTGCCAAAAGTCAAAA 208

QY 404 TGGAAATCGTTGTTAAGAAAGACCAAGTGAATCTGTAATCAACAAATAATTGAAGGAG 463
```

Db 209 TCGAGTGGTGGCGGATGATCCGTCGACGCGGATTCGAGTGAATGTCGAGGTGG 268
 QY 464 CAAGGACAGGAGAGATGCGTATGCGACAGATTTTGTGCTGCTGTCAGATGTCATAA 523
 Db 269 CGCGTTCCGGGCAAAATCGCGACGCGCAAGATTTTGTGCTGCGGTCGAGGAGGCAATCC 328
 QY 524 GAGTTAGGACAGGTGAGGTGGGGAAGCAG 556
 Db 329 GTATCCGACGCGGCAACGTTCCGACGCGGCGG 361

RESULT 15

AAAB1472
 ID AAA81472 standard; DNA; 16526 BP.
 AC AAA81472;
 XX
 DT 04-DEC-2000 (first entry)
 XX
 DE N. meningitidis partial DNA sequence gnm_20 SEQ ID NO:20.
 XX
 KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
 KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
 KW Meningococcus B; MenB; ds.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO20002243C-A2.
 PD 20-APR-2000.
 XX
 PF 08-OCT-1999; 99WO-US023573.
 XX
 PR 09-OCT-1998; 98US-0103794P.
 PR 30-APR-1999; 99US-0132069P.
 XX
 PA (CHIR) CHIRON CORP.
 XX
 PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
 PI Maignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
 PI Rappuoli R, Pizza M;
 DR WPI; 2000-318079/27.

Isolated nucleotide sequences of Neisseria meningitidis which can be used in the diagnosis and treatment of N. meningitidis infection and other Neisseria infections, for example, N. gonorrhoea.

Claim 7; Page 466-471; 1760pp; English.

The present invention describes methods of obtaining immunogenic proteins from Neisseria genomic sequences. AAA81453 to AAA82414 represent specifically claimed Neisseria meningitidis genomic DNA sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent Neisseria DNA sequences and their corresponding proteins; AAA81254 to AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the isolation of Neisseria meningitidis DNA sequences; and AAA81322 to AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF sequences, which are all used in the exemplification of the present invention. The nucleic acid sequences, protein sequences, and antibodies against them, can be used in the manufacture of a composition. The composition can be used as a medicament for in the manufacture of a medicament for treating, preventing or diagnosing infection due to Neisseria bacteria. For example, some of the identified proteins could be components of vaccines against Meningococcus B; against all serotypes; and/or against all pathogenic Neisseriae. Identification of sequences from the bacterium will also facilitate production of biological probes, particularly organism-specific probes. Attempts to make efficacious Meningococcus B vaccines have failed mainly due to antigen tolerance. Multivalent vaccines have also been tried but none have successfully overcome antigenic variability. The provision of further, complete sequences may provide an opportunity to identify secreted or surface exposed proteins

CC that may be presumed targets for the immune system and which are not
 CC antigenically variable or at least more conserved than other more
 CC variable regions

XX Sequence 16526 BP; 3183 A; 4006 C; 5067 G; 4268 T; 0 U; 2 Other;

Query Match 12.4%; Score 73; DB 3; Length 16526;

Best Local Similarity 55.3%; Pred. No. 4.4e-11;
 Matches 184; Conservative 0; Mismatches 145; Indels 4; Gaps 2;

QY 224 ACAAGTGGGAAGCAATTGTCAGACCAATGAGAAATCCAGCAAGTTTCATCGGCTTTACTGA 283

Db 13603 AAAAAATCGAGGCGATTGTCAAAACCGTTCAAACTCGACACGTGCGGAGGCGCTTGACCG 13662

QY 284 AAATCGGATTCGAGGTGTACTGTCTTCTGATGTGAGAGGTTTGGTGCACAGGAGGTT 343

Db 13663 AAATCGGCATTACGGGCATGACCGTCAAGCGGTCAAGGGTTCGGCAGGCAGAGGGGC 13722

QY 344 CTACCGAGAGACACGCTGCTCTGAGTTCTCGAAGACAAAATTGTTGCTAAAGTTAAGA 403

Db 13723 ATACGGAATCTATCGCGCGGGATACGCCGTGAT---TTCCTGCCCAATCAAAA 13779

QY 404 TCGAAATCGTTGTTAAGAAAGACCAAGTCGATCTGTAATCAACACAATTAATTGAAGGAG 463

Db 13780 TCGAGCTGTTGTCGGGATGATGCTGTGACGCGCGGATTCACGTGAT-TTCGAGGTGG 13838

QY 464 CAAGGACAGGAGAGATTGTTGATGCGCAAGATTTTGTGCTGCTGTCAGATGTCATAA 523

Db 13839 CGCGTTCCGGGAAAAATCGCGCAGCGCAAGATTTTGTGCTGCGGTTGAGAGGCAATCC 13898

QY 524 GAGTTAGGACAGGTGAGCGTGGGGAAGCAG 556

Db 13899 GTATCCGACGCGGCAACGTTCCGACGCGGCGAG 13931

Search completed: May 27, 2004, 19:59:42

Job time : 207.069 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2004, 13:35:34 ; Search time 1755.75 Seconds
(without alignments)
14663.718 Million cell updates/sec

Title: US-09-756-541-16
Perfect score: 594
Sequence: 1 CGGCACGAGGCTACTGCGAA.....CTGACATGAGTACTTCTGCT 594

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 1 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb ba.*
- 2: gb b-g.*
- 3: gb in.*
- 4: gb om.*
- 5: gb ov.*
- 6: gb pat.*
- 7: gb ph.*
- 8: gb pl.*
- 9: gb pr.*
- 10: gb ro.*
- 11: gb sts.*
- 12: gb sy.*
- 13: gb un.*
- 14: gb vi.*
- 15: em ba.*
- 16: em fun.*
- 17: em hum.*
- 18: em in.*
- 19: em nu.*
- 20: em om.*
- 21: em or.*
- 22: em ov.*
- 23: em pat.*
- 24: em ph.*
- 25: em pl.*
- 26: em ro.*
- 27: em sts.*
- 28: em un.*
- 29: em vi.*
- 30: em htg hum.*
- 31: em htg inv.*
- 32: em htg other.*
- 33: em htg mus.*
- 34: em htg pin.*
- 35: em htg rod.*
- 36: em htg mam.*
- 37: em htg vrt.*
- 38: em sy.*
- 39: em htgo hum.*
- 40: em htgo mus.*
- 41: em htgo other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	594	100.0	594	6	AR125592	AR125592 Sequence
2	594	100.0	897	6	AR125590	AR125590 Sequence
3	592	99.7	840	8	AF095454	AF095454 Ricinus c
4	300.2	50.5	796	8	AY027892	AY027892 Medicago
5	285	48.0	588	6	AR125591	AR125591 Sequence
6	285	48.0	591	6	AX507438	AX507438 Sequence
7	285	48.0	591	8	BT005209	BT005209 Arabidops
8	285	48.0	817	6	AR125589	AR125589 Sequence
9	285	48.0	844	8	AF095455	AF095455 Arabidops
10	275.2	46.3	947	8	AY442185	AY442185 Lycopersi
11	216.2	36.4	902	8	AK068407	AK068407 Oryza sat
12	216.2	36.4	902	8	AK099152	AK099152 Oryza sat
13	215.4	36.3	1085	8	PII489604	PII489604 Pinus pin
14	107	18.0	686	1	AF017419	AF017419 Nostoc pu
15	103.2	17.4	2493	1	ASF251822	ASF251822 Anabaena
16	103.2	17.4	341880	1	AP003589	AP003589 Nostoc sp
17	99	16.7	384	1	FDGLNBPRT	FDGLNBPRT X97327 F.diplosiph
18	98.8	16.6	191028	8	PPU38804	PPU38804 Porphyra pu
19	96.4	16.2	11978	1	U67574	U67574 Methanococc
20	96.4	16.2	110000	6	AR271569_01	AR271569_01 Continuation (2 of
21	96.2	16.2	10338	1	U67464	U67464 Methanococc
22	96.2	16.2	110000	6	AR271569_14	AR271569_14 Continuation (15 o
23	95.4	16.1	383	1	SYOGLNB	SYOGLNB M62447 Synecchococ
24	95.4	16.1	4348	1	AF079137	AF079137 Synecchoc
25	94.4	15.9	452	1	SSGLNBP	SSGLNBP X97496 Synecchocyst
26	94.4	15.9	130001	1	D90915	D90915 Synecchocyst
27	92	15.5	300143	1	AE017165	AE017165 Prochloro
28	90.4	15.2	164321	8	AF022186	AF022186 Cyanidium
29	90.4	15.2	299350	1	AP005370	AP005370 Thermosyn
30	88.4	14.9	9973	1	AE004295	AE004295 Vibrio ch
31	88.2	14.8	1622	1	AF120107	AF120107 Synecchoc
32	86.8	14.6	30045	1	AE016803	AE016803 Vibrio vu
33	86.8	14.6	302320	1	AP005081	AP005081 Vibrio pa
34	86.4	14.5	262202	1	EX572094	EX572094 Prochloro
35	85.2	14.3	247950	1	AP005340	AP005340 Vibrio vu
36	84.8	14.3	96109	6	AR408756	AR408756 Sequence
37	84.8	14.3	96109	6	AX067460	AX067460 Sequence
38	84	14.1	11371	1	AE006236	AE006236 Pasteurel
39	83.6	14.1	375	6	AR318761	AR318761 Sequence
40	83.6	14.1	298950	1	AP004597	AP004597 Oceanobac
41	83.4	14.0	9877	1	AE002548	AE002548 Neisseria
42	83.2	14.0	300880	1	AE016917	AE016917 Chromobac
43	82.2	13.8	300272	1	AE017213	AE017213 Geobacter
44	82	13.8	346792	1	EX571658	EX571658 Wolinella
45	81.8	13.8	349061	1	NMA222491	NMA222491 Neisseria

ALIGNMENTS

RESULT 1
AR125592
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES

AR125592
Sequence 16 from patent US 6177275.
AR125592
AR125592.1 GI:14111654
Unknown.
Unknown.
Unclassified.
1 (bases 1 to 594)
Coruzzi,G.M., Lam,H.-M. and Hsieh,M.-H.
Plant nitrogen regulatory P-II Genes
Patent: US 6177275-A 16 23-JAN-2001;
Location/Qualifiers

linear PAT 16-MAY-2001

QY	1	CGGCACAGGCTACTCGGAACCTGGGCTTGGCTCACTCTCTTCATTCTAATAACATCAAG	60	
Db	50	CGGCACAGGCTACTCGGAACCTGGGCTTGGCTCACTCTCTTCATTCTAATAACATCAAG	109	
QY	61	AAAGAAATTCCTGTTTGTGATTTCAGTTTGTGTTTGTCCAGAGCTTAGACAATCTCGGTTT	120	
Db	110	AAAGAAATTCCTGTTTGTGATTTCAGTTTGTGTTTGTCCAGAGCTTAGACAATCTCGGTTT	169	
QY	121	TCTCACTTTTAACACCGCGGTCAAGCGCTTAAGATATGCCCCGTCGTCCTGTGATTAAT	180	
Db	170	TCTCACTTTTAACACCGCGGTCAAGCGCTTAAGATATGCCCCGTCGTCCTGTGATTAAT	229	
QY	181	GCCCAAGCTCGCTGACTACATTCCTGATGCTAAATTTCTACAAAGTGAAGCAATTCCTC	240	
Db	230	GCCCAAGCTCGCTGACTACATTCCTGATGCTAAATTTCTACAAAGTGAAGCAATTCCTC	289	
QY	241	AGGCCCTGGCGAGTCTCGCAAGTTTCCCTCGGCTTTGCTAAAAATTTGGTATTCGAGGTGT	300	
Db	290	AGGCCCTGGCGAGTCTCGCAAGTTTCCCTCGGCTTTGCTAAAAATTTGGTATTCGAGGTGT	349	
QY	301	ACTGTTTCTGATGTTCCAGGTTTTCGGTGCTCAAGTGCTTCAACTGAGAGGCGAGGCGGC	360	
Db	350	ACTGTTTCTGATGTTCCAGGTTTTCGGTGCTCAAGTGCTTCAACTGAGAGGCGAGGCGGC	409	
QY	361	TCAGAAATTTCTGAAGACAAAGTTTCTGCTAAAGTTTAAGATCGAGATCGTGTTAGCAA	420	
Db	410	TCAGAAATTTCTGAAGACAAAGTTTCTGCTAAAGTTTAAGATCGAGATCGTGTTAGCAA	469	
QY	421	GACCAGGTTGAGGATGTTATAGAAAAATCAITGAGAGGCGCAAGAACTCGAGAGATTGGA	480	
Db	470	GACCAGGTTGAGGATGTTATAGAAAAATCAITGAGAGGCGCAAGAACTCGAGAGATTGGA	529	
QY	481	CAOCCCAAGATTTCTTGCTGCTGTTTCAGATGTTAATAAGATCCGCACTCGTGAGCGG	540	
Db	530	CAOCCCAAGATTTCTTGCTGCTGTTTCAGATGTTAATAAGATCCGCACTCGTGAGCGG	589	
QY	541	GGTGATAGGCTGACAGATGACAGAGGCGCATCTGACATGATGACTTCTTGCT	594	
Db	590	GGTGATAGGCTGACAGATGACAGAGGCGCATCTGACATGATGACTTCTTGCT	643	
RESULT 3	AF095454	840 bp	linear	PLN 19-NOV-1998
LOCUS	Ricinus communis PII protein mRNA, partial cds.			
DEFINITION	AF095454			
ACCESSION	AF095454.1	GI:3885940		
VERSION				
KEYWORDS				
SOURCE	Ricinus communis (castor bean)			
ORGANISM	Ricinus communis			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta			
AUTHORS	Spematophyta; Magnoliophyta; eudicotyledons; core eudicots;			
TITLE	rosids; eurosids I; Malpighiales; Euphorbiaceae; Acalyphoideae;			
	Acalyphaeae; Ricinus.			
	1 (bases 1 to 840)			
	Hsieh, M.H., Lam, H.M., van de Loo, F.J. and Coruzzi, G.			
	A PII-like protein in Arabidopsis: putative role in nitrogen			
	sensing			
	Proc. Natl. Acad. Sci. U.S.A. 95 (23), 13965-13970 (1998)			
JOURNAL	99030678			
MEDLINE	9811909			
PUBMED				
REFERENCE	2 (bases 1 to 840)			
AUTHORS	Hsieh, M.H., Lam, H.M., van de Loo, F.J. and Coruzzi, G.M.			
TITLE	Direct Submission			
JOURNAL	Submitted (29-SEP-1998) Biology, New York University, 100			
	Washington Square East, New York, NY 10003, USA			
FEATURES	Location/Qualifiers			
source	1..840			
	/organism="Ricinus communis"			
	/mol_type="mRNA"			
	/db_xref="taxon:3988"			
	<1..595			
CDS				

/function="putative role in nitrogen sensing"

/note="GLB1"
/codon_start=2
/product="PII protein"
/protein_id="AAC78332.1"
/db_xref="GI:3885941"
/translation="HEATKILGLLPLHNNIKKEFPVDFESLPCBLHRSFHFNT
AVKRYAPVPVINAQSSDYPDAKFYKVEAILRPWRVQVSSALLKIGIRGVTVS
DVRFQAGGSGTERQGGSEFSEDFAKVMEIIVSKDQVEDVIEKIIEBARTGEIGD
GKIFLPVSDVIRVIRTCBERGDKAERTMGSRDMSTSA"

ORIGIN

Query Match 99.7%; Score 592; DB 8; Length 840;
Best Local Similarity 100.0%; Pred. No. 2.5e-153;
Matches 592; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GCACGAGGCTACTCGAAACTGGGCTTCTCACTCTCTTCTTCTAATAACATCAAGAA 62
Db 1 GCACGAGGCTACTCGAAACTGGGCTTCTCACTCTCTTCTTCTAATAACATCAAGAA 60
QY 63 AGAATCCCTGTTTTGATTTTCAGTTTGTTCAGAGCTTAGACATTCGGTTTTTC 122
Db 61 AGAATCCCTGTTTTGATTTTCAGTTTGTTCAGAGCTTAGACATTCGGTTTTTC 120
QY 123 TCACATTAAACCCGGGTCAAGCGGTAAAGATATGCCCGGCTGTTCTGTGATTAATGC 182
Db 121 TCACATTAAACCCGGGTCAAGCGGTAAAGATATGCCCGGCTGTTCTGTGATTAATGC 180
QY 183 CCAAGCTCGCTGACTACATTCCTGATGCTAAATTTTACAAAGTGAAGCAATTCCTCAG 242
Db 181 CCAAGCTCGCTGACTACATTCCTGATGCTAAATTTTACAAAGTGAAGCAATTCCTCAG 240
QY 243 GCCCTGGGAGTCTCGCAAGTTTCTCGGCTTGTCTTAAATTTGTTATTCAGAGTTTAC 302
Db 241 GCCCTGGGAGTCTCGCAAGTTTCTCGGCTTGTCTTAAATTTGTTATTCAGAGTTTAC 300
QY 303 TGTTCATGATTCGAGTTTGTGCTCAAGTGGTTCAACTGAGAGCGGGGGCTC 362
Db 301 TGTTCATGATTCGAGTTTGTGCTCAAGTGGTTCAACTGAGAGCGGGGGCTC 360
QY 363 AGAATTTCTCAAGACAAGTTTGTCTTAAAGTTAAGATGAGATCGTGGTTAGCAAGA 422
Db 361 AGAATTTCTCAAGACAAGTTTGTCTTAAAGTTAAGATGAGATCGTGGTTAGCAAGA 420
QY 423 CCAGTTGAGGATGTTATAGAAAATCATTGAGGAGCAAGAACTGAGAGATTGGAGA 482
Db 421 CCAGTTGAGGATGTTATAGAAAATCATTGAGGAGCAAGAACTGAGAGATTGGAGA 480
QY 483 CGGCAAGATTTTCTGCTGCTTTCAGATGTAATAGATCCGCACTGGTGAGCGGG 542
Db 481 CGGCAAGATTTTCTGCTGCTTTCAGATGTAATAGATCCGCACTGGTGAGCGGG 540
QY 543 TGATAAGGCTGAGAGGATGACAGGAGCGGATCTGACATGACTTCTGCT 594
Db 541 TGATAAGGCTGAGAGGATGACAGGAGCGGATCTGACATGACTTCTGCT 592

RESULT 4
AY027892
LOCUS Medicago sativa PII protein (GLNB) mRNA, complete cds.
DEFINITION Medicago sativa PII protein (GLNB) mRNA, complete cds.
ACCESSION AY027892
VERSION AY027892.1 GI:13277514
KEYWORDS
SOURCE Medicago sativa
ORGANISM Medicago sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE 1 (bases 1 to 796)
AUTHORS Garcia-Ibáñez, D. and Sengupta-Gopalan, C.
TITLE Characterization of PII (GLNB) in alfalfa

JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

FEATURES

source

Unpublished
2 (bases 1 to 796)
Garcia-Ibáñez, D. and Sengupta-Gopalan, C.
Direct Submission
Submitted (22-FEB-2001) Molecular Biology Program, New Mexico State
University, Corner of Knox and College, Las Cruces, NM 88003, USA
Location/Qualifiers
1..796
/organism="Medicago sativa"
/mol_type="mRNA"
/db_xref="taxon:3879"
1..796
/gene="GLNB"
/gene="GLNB"
/codon_start=1
/product="PII protein"
/protein_id="AAK16221.1"
/db_xref="GI:13277515"
/translation="WTLLAKPVENGLNPINETOIPSSFSVIRKPCDSSHRNVVL
RSGNASVLPKIRAKNPEDYVPESKFYVEAILRPWRIPQVSSGLLKMIRGVTVDV
KGFQAGGSKERQGGSEFSEDFAKVMEIIVRKDQVEAVINKIMETARTGEIGDK
IELIPVSDVIRITRGERGEQARMAGGLTDALYV"

ORIGIN

Query Match 50.5%; Score 300.2; DB 8; Length 796;
Best Local Similarity 72.0%; Pred. No. 2.8e-72;
Matches 408; Conservative 0; Mismatches 153; Indels 6; Gaps 1;

QY 13 ACTCGAACTGGGCTGCTCACTCTCTTCAATTAATAACATCAAGAAAGAAATTCCT 72
Db 10 ATTGGAACCGACGTGTTCAACGGTTTGAATTTAATATCAATCAAACTCAAATTCCT 69
QY 73 GTTTTGAATTTCAAGTTTGTTCAGAGCTTAGACATTCGGTTTCTACITTAAC 132
Db 70 TTTTCAAGCTTCAGGCTCATTCGCAAGCGTTTCGAGATTC-----TTCTCATCGCAAT 123
QY 133 ACCGGGTCAAGCGGTAAAGATATGCCCGGCTGTTCTGTGATTAATGCCCAAACTCG 192
Db 124 GTGGTCTTAAGAAGCAACGGAATGCTTCGGTTCTTCCCAAAATTAGAGCTCAAAACCTT 183
QY 193 CTGACTACATTCCTGATGCTAAATTTCTACAAAGTGAAGCAATTCAGAGGCTCGCGA 252
Db 184 CTGACTATGTTCTGATTCAGATTCAGTTTACAAAGTTGAAGCCATTCAGGCCATCGAGA 243
QY 253 GTCTCGCAAGTTTCTCGGCTTGTGCTAAATTTGTTATTCAGAGTGTACTTCTGAT 312
Db 244 ATCCCTCAGGTTTCTTCGGGTTGTTGAAATGGGAATTCGTGTGTCTACTGTATCTGAT 303
QY 313 GTTCGAGTTTGTGCTCAAGTGTGTTCAACTGAGAGCGAGGCGGCTCAGAAATTTCT 372
Db 304 GTCAAGGGTTTGTGCTCAGGTTGCTCAAAAGAGAGCGAGGAGGCTCTGAATTTCT 363
QY 373 GAAGACAAGTTTGTGCTAAAGTTAAGATGAGATGCTGTTAGCAAAAGACCAAGTTGAG 432
Db 364 GAAGACAATTTTGTGCCAAAGTTAAATGGAAATAGTGTGTGAGAAAGACCAAGTTGAG 423
QY 433 GATGTTATAGAAAAATCATTGAGAGCGCAAGCACTGGAGAGATTGGAGAGCGCAAGATT 492
Db 424 GCGATGATAGAAAAATTTATGAGAGCGCAAGCACTGGGAGAGATTGGATGAGCAAAAT 483
QY 493 TTTCTGCTGCTGTTTTCAGATGTTAATAGATCCGCACTGTTGAGCGGGGTGATAGGCT 552
Db 484 TTTCTGATCCCTGTTATCTGATGTTAATAGATCCGCACTGTTGAGCGGGGTGATAGGCT 543
QY 553 GAGAGGATGACAGGAGGGGATCTGAC 579
Db 544 GAGAGGATGCTGGGGAGCTAATGAC 570

RESULT 5
AR125591
LOCUS

AR125591 588 bp DNA linear PAT 16-MAY-2001

DEFINITION	Sequence 15 from patent US 6172725.			
ACCESSION	ARI25591	GI:14111653		
VERSION	ARI25591.1			
KEYWORDS	Unknown.			
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	1 (bases 1 to 588)			
AUTHORS	Coruzzi, G.M., Lam, H.-M. and Hsieh, M.-H.			
TITLE	Plant nitrogen regulatory P-PII genes			
JOURNAL	Patent: US 6172725-A 15 23-JAN-2001;			
FEATURES	Location/Qualifiers			
source	1..588			
	/organism="unknown"			
	/mol_type="unassigned DNA"			
ORIGIN				
Query Match	48.0%;	Score 285;	DB 6;	Length 588;
Best Local Similarity	70.3%;	Pred. No. 4.5e-68;		
Matches 397;	Conservative 0;	Mismatches 165;	Indels 3;	Gaps 17;
Qy	7	GAGGCTACTGCGAAACTGGGCTTGCTCACTCCTCTTCATTCTTAATAACATCAAGAAAGA	66	
Db	7	CGGTCATGACGAAACCCTCTCAATAACTTCTCTCGGTTTCTATTCTGATCGAAGAAC	66	
Qy	67	TTCCCTGTTTTTGATTTTCAGTTTGTGTTGTCAGAGCTTAGACATCTCGGTTTTCTCAC	126	
Db	67	ATTGCTTTCTCTGATGTCATTTCGATTGTTGTCGATTTCAGACATCTCCGACCATCTTGC	126	
Qy	127	TTTAAACACCGCGGTCAAGCGGTAAGATATGCC---CCGTCGTTTCTGTGATTAATGCC	183	
Db	127	CTCGATTGGTTCACAAAGTCACGAGTAATAACAGTCGTGTTTACCTGCTGTAGTCC	186	
Qy	184	CAAAAGTCGCTGACTACATTCCTGATGCTAAAATTTCTACAAGTGAAGCAATTTCTCAGG	243	
Db	187	CAAAATATCTTCGTATTAATATCCAGACTCGAAAATTTTACAAGTGAAGCAATTTCTCAGA	246	
Qy	244	CCCTGGCGAGTCTCGCAAGTTTCTCCGCTTTGCTTAAAAATGGSPATTCGAGGTGTACT	303	
Db	247	CCATGGAGAAATCCAGCAAGTTTCATCGCGCTTTACTGAAAAATCGGGATTCGAGGTGTACT	306	
Qy	304	GTTTCTGATGTTCCGAGGTTTTTGGTGTCTCAAGTGTGTTCAAATCTGAGAGCGAGGCGGCTCA	363	
Db	307	GTTTCTGATGTGAGAGGGTTTGGTGCAACAGGAGTTCTACCGAGAGACACGGTGGCTCT	366	
Qy	364	GAATTTCTGAAGACAATTTGTTGCTAAAGTTAAGATGGAGATCGTGGTAGCAAGAC	423	
Db	367	GAGTTCTCGGAAGACAAATTTGTTGCTAAAGTTAAGATCGAATCGTTTAAAGAAAGAC	426	
Qy	424	CAGGTTGAGGATGTTATAGAAAAATCATTCAGGAGGCAAGAACTGGAGAGATTGCGAC	483	
Db	427	CAAGTGGAAATCTGTAATCAACACAATAATTGAGGACAGGACAGGAGAGATTGGTGAT	486	
Qy	484	GGCAAGATTTTCTTGCTGCCTGTTTTCAGATGTAATAAGAGTCGCGACCTGGTGAGCGGGT	543	
Db	487	GGCAAGATTTTGTGTTTTCGCTGTGTCAGATGTCTATAAGAGTTAGGACAGGTGAGCGTGG	546	
Qy	544	GATTAAGGCTGAGGATGACAGGAG	568	
Db	547	GAGAAAGCAGAGAAATGACTGGTG	571	

RESULT	6
LOCUS	AX507438
DEFINITION	Sequence 2133 from Patent WO0216655.
ACCESSION	AX507438
VERSION	AX507438.1
KEYWORDS	. GI:23388675
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
	linear PAT 27-SEP-2002

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1
 REFERENCE
 AUTHORS Harper, J.F., Kreps, J., Wang, X. and Zhu, T.
 TITLE Stress-regulated genes of plants, transgenic plants containing same, and methods of use
 JOURNAL Patent: WO 0216655-A 2133 28-FEB-2002;
 The Scripps Research Institute (US) ; Syngenta Participations AG (CH)

FEATURES
 source Location/Qualifiers
 1. 591
 /organism="Arabidopsis thaliana"
 /mol_type="unassigned DNA"
 /db_xref="taxon:3702"

ORIGIN
 Query Match 48.0%; Score 285; DB 6; Length 591;
 Best Local Similarity 70.3%; Pred. No. 4.5e-68;
 Matches 397; Conservative 0; Mismatches 165; Indels 3; Gaps 1

Qy 7 GAGGCTACTCGGAACCTGGGCTTGCTCACTCCTCTTCATCTTAATAACATCAAGAAAGAA 66
 Db 7 GCGTCAATGACGAACCCATCTCAATACATCTCTCGGTTCTATTCTGATCGAAGAAGC 66
 Qy 67 TTCCTGTTTGTGATTTTCAGTTTGTGTTGTCAGAGCTTAGACATCTCGGTTTTCACAC 126
 Db 67 ATTGCTTTCTCGATGTCATTTTCGATTTGTTCTGGATTTCAGACATTCGCGACATCTTGC 126
 Qy 127 TTTAAACACCGGGTCAAGCGGTAAGATATGCC---CCGTCGTTCTCTGTGATTAAATGCC 183
 Db 127 CTCGATTTGGTCAAAAGTCAACAGTCAACGAGTAATACAGTCGTGTTTACCTGCTGTTAGTCC 186
 Qy 184 CAAAGCTCGCGTCACTACATCTCTGATGCTAAATTTCTACAAAGTGGAAAGCAATTCCTAGG 243
 Db 187 CAAATATCTTCTGATTATATATCCAGACTCGAAATTTTACAAAGTGGAAAGCAATTCCTAGA 246
 Qy 244 CCCTGGCGAGTCTCGAAGTTTCCTCGGCTTGTCTAAAAATGGTATTTCGAGGTCTACT 303
 Db 247 CCATGAGAAATCCAGCAAGTTTCATCGGCTTTACTCTGAAAATGCGGATTCGAGGTGTACT 306
 Qy 304 GTTTCATGTTCCGAGGTTTTGGTGTCTCAAGTGTCTCAACTGAGAGGCGAGGCGGCTCA 363
 Db 307 GTTTCATGAGAGGGTTTTGGTGTCACAAGAGGTTCTACGAGAGACACGTTGGCTCT 366
 Qy 364 GAATTTCTGAAGCAAGTTTGTGCTAAAGTTAAGATGGAGATCGTGTTAGCAAGAC 423
 Db 367 GAGTTCGGAAGCAAAATTTGTTGCTAAAGTTAAGATGGAAATCGTTGTTAAGAAAGAC 426
 Qy 424 CAGGTTGAGGATGTTATAGAAAAAATCATTTGAGGAGGCAAGACTGGAGAGATTGGAGC 483
 Db 427 CAAGTGGAAATCTGTAATCAACACAATATTTGAAGGAGCAAGGACAGGAGAGATTGGTGT 486
 Qy 484 GCGAAGATTTTCTGTCGCCGTTTCAGATCTTAATAGAGTCCGCACTGGTCAGCGGGGT 543
 Db 487 GCGAAGATTTTGTGTTTGGCTGTGTGTCAGATGTCATAGAGTTAGGACAGGTGAGCGTGG 546
 Qy 544 GATAAGGCTCAGAGGATGACAGGAG 568
 Db 547 GAGAAAGCAGAGAGATGACTGGTG 571

RESULT 7
 BT005209
 LOCUS Arabidopsis thaliana At4g01900 mRNA, complete cds.
 DEFINITION
 ACCESSION BT005209
 VERSION BT005209.1
 KEYWORDS GI:28950698
 SOURCE FLI CDNA.
 ORGANISM Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE
 1 (bases 1 to 591)

	Db	67	ATTGCTTTCCTGATGCATTTTCGATTGTCTGGGATTCAGACATTCGCCAGCATCTTGC	126
	Qy	127	TTTAAACCCCGCGGTCAAGCGCGTAAGATATGCC--CCGCTCGTTCTCTGATTAATGCC	183
	Db	127	CTCGATTTGGTCACAAAGTCACCGAGTAATAACAGTCGTGTTTTACTCTGTCTGTTAGTGCC	186
	Qy	184	CAGAAGCTCGCTCAGCTACATCTCTGATGCTTAATCTCACAAAGTGGGAACAATCTTCAGG	243
	Db	187	CAAATATCTTCTGATTATATTCAGACTCGAAATTTTACAAGGTGGAAGCAATTTGTTCAGA	246
	Qy	244	CCCTGGCGAGTCTCGCAAGTTTCTCGCTTTCCTAATAAATTCGTTATTCAGAGTGTACT	303
	Db	247	CNATGGAGAAATCAGCAAGTTTCATCGGCTTACTGAAAATTCGGGATTCAGAGTGTACT	306
	Qy	304	GTTTCTGATGTTTCGAGGTTTTTGCTGCTCAAGTGGTTTCAACTGAGAGCGCAGGCGCTCA	363
	Db	307	GTTTCTGATGTGAGAGGGTTTTGGTGCACAAGGAGGTTCTACCAGAGACACCGGTGGCTCT	366
	Qy	364	GAAATTTCTGAAGACAAGTTTGTTCCTAAAGTTTAAGTAGGAGATCGTGGTTAGCAAGAC	423
	Db	367	GAGTTCTCGGAAGACAAATTTGTTGCTTAAAGTTTAAAGTGGAAATCGTTGTTTAAAGAAGAC	426
	Qy	424	CAGTTTGAGGATCTTATAGAAAAATCATTTGAGGCGCAGAAGCTGCAGAGATTTGGAGAC	483
	Db	427	CAAGTGGATCTGTAAATCAACACAAATTTGAAGGACAGGACAGGAGAGATTTGGTAT	486
	Qy	484	GGCAAGATTTTCTTGCTGCTCTTTTCAGATGTAATAAGAGTCCGCACTCGGTGAGCGGGT	543
	Db	487	GGCAAGATTTTGTGTTGCTGTGTGAGATGTCATTAAGAGTTAGGACAGGTGAGCGTGGG	546
	Qy	544	GATAAGCGTCGAGAGATGACAGGAG	568
	Db	547	GAGAAAGCAGAGAAGATGACTGGTG	571
RESULT	8			
ARL125589				
LOCUS		ARL125589	817 bp	DNA linear PAT 16-MAY-2001
DEFINITION		Sequence 13 from patent US 6177275.		
ACCESSION		ARL125589		
VERSION		ARL125589.1 GI:14111651		
KEYWORDS				
SOURCE		Unknown.		
ORGANISM		Unclassified.		
REFERENCE		1 (bases 1 to 817)		
AUTHORS		Coruzzi,G.M., Lam,H.-M. and Hsieh,M.-H.		
TITLE		Plant nitrogen regulatory P-Pil genes		
JOURNAL		Patent: US 6177275-A 13 23-JAN-2001;		
FEATURES		Location/Qualifiers		
source		1..817		
		/organism="unknown"		
ORIGIN		/mol_type="unassigned DNA"		
Query Match		48.0%; Score 285; DB 6; Length 817;		
Best Local Similarity		70.3%; Pred. No. 4.7e-68;		
Matches		Conservative 0; Mismatches 165; Indels 3; Gaps 1;		
Qy		7	GAGGCTACTGCGAAACTGGGCTGCTCACTCCCTCTTCAATTCCTAATAACATCAAGAAAGAA	66
Db		39	GCCTCAATGACGAAACCCATCTCAATAACTTCTCTCGGTTTCTATTCTGATCGAAGAAGAC	98
Qy		67	TTCCCTGTTTTTCATTTCAGTTTGTGTTGTCAGAGCTTAGACATTCCTGGTTTTTCTCAC	126
Db		99	ATTGCTTTTCTGATTGCAATTCGATTGTGTTGTTGATTCAGACATTCGCCAACCTCTTGC	158
Qy		127	TTTAAACCCCGCGGTCAAGCGCGTAAGATATGCC--CCGCTCGTTCTCTGATTAATSCC	183
Db		159	CTCGATTTGGTCACAAGTCCAGGATTAACAGTCGTGTTTTTACCTGCTGTTAGTGCC	218
Qy		184	CAGAAGCTCGCTCAGCTACATTCCTGATGCTAAATTTCTAAGAAGTGAAGCAATTCCTCAGG	243

```

Db      219 CAAATATCTTCTGATTATATCCAGACTCGAATTTTCAAGGTGAGCAATTGTGAGA 278
QY      244 CCCTGCGAGTCTCGCAAGTTTCTCGGCTTCTCTAAATTTGTTAGTATTCAGGTGTTACT 303
Db      279 CCATGAGAAATCAGCAAGTTTCTCGGCTTCTCTAAATTTGTTAGTATTCAGGTGTTACT 338
QY      304 GTTCTGATGTTTCGAGTTTGTGCTCAAGTGGTTCACCTGAGAGGCGGCGGTCA 363
Db      339 GTTCTGATGTGAGAGGTTTGTGCAAGGAGGTTCTACCGAGAGACACGGTGGCTCT 398
QY      364 GAATTTTCTGAACAAGTTTGTGCTAAAGTTAAAGATGAGATCGTGGTTAGCAAGAC 423
Db      399 GAGTTCTCGAAGACAAATTTGTCTAAAGTTAAGTGAATCGTTGTTAGAAGAC 458
QY      424 CAGGTTGAGGATGTTATAGAAAAATCATTTGAGGAGGCAAGAACTGAGAGATGAGAGAC 483
Db      459 CAAGTGGATCTGTAATCAACACATAATTTGAAGGAGCAAGGACAGAGAGATGAGTGT 518
QY      484 GCGAAGATTTCTGCTGCTGCTGTTTCAGATGTAATAGAGTCCGCACTGGTGAAGCGGGT 543
Db      519 GCGAAGATTTTGTGCTGCTGCTGTTTCAGATGTCATAGAGTTAGGACAGGTGAGCGTGG 578
QY      544 GATAAGCTGAGAGGATGACAGGAG 568
Db      579 GAGAAAGCAGAGAGATGACTGTG 603

RESULT 9
AF095455 844 bp mRNA linear PLN 19-NOV-1998
Arabidopsis thaliana PII protein mRNA, complete cds.
DEFINITION
ACCESSION AF095455
VERSION AF095455.1 GI:3885942
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
Hsieh,M.H., Lam,H.M., van de Loo,F.J. and Coruzzi,G.
A PII-like protein in Arabidopsis: putative role in nitrogen
sensing
Proc. Natl. Acad. Sci. U.S.A. 95 (23), 13965-13970 (1998)
MEDLINE 99030678
PubMed 9811909
REFERENCE 2 (bases 1 to 844)
AUTHORS Hsieh,M.H., Lam,H.M., van de Loo,F.J. and Coruzzi,G.M.
Direct Submission
JOURNAL Submitted (29-SEP-1998) Biology, New York University, 100
Washington Square East, New York, NY 10003, USA
FEATURES
Location/Qualifiers
1..844
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chrosome="1V"
/map="10.8 cm"
31..621
/function="putative role in nitrogen sensing"
/notes="GLB1"
/codon_start=1
/product="PII protein"
/protein_id="AAC7833.1"
/db_xref="GI:3885943"
/translation="MAAMTKPISITSLGFSYDRKNIAFSDICISCSGFRHSRPSCLD
LVTKSPNNRSLVPLVSAQISSDYIPDSKFKVEAVRPNRIQOVSSALLKIGRGVT
VSDVRGFGAGSGSFRHGSEFSEDFKFAKVMELIVVKDQVESVINTIIGARTGEI
GGKIFVLVSDVIRVKTGRGKAEKMGDMLSPS"

CDS
Query Match 48.0%; Score 285; DB 8; Length 844;
Best Local Similarity 70.3%; Pred. No. 4.7e-68;

ORIGIN

```

```

Matches 397; Conservative 0; Mismatches 165; Indels 3; Gaps 1;
QY      7 GAGGCTACTGCGAAACTGGGCTTGCTCAGCTCTCTTCTAATTAACATCAAGAAAGAA 66
Db      37 GCGTCAATGACGAACCCCATCTCAATAACTTCTCTCGGTTTCTTATCTGATCGAAGAAC 96
QY      67 TTCCCTGTTTTCGATTCAGTTTGTGTCAGAGCTTAGACATTCCTGGTTTCTCTCAC 126
Db      97 ATTGCTTCTCTGATTCGATTCGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 156
QY      127 TTTAACACCGCGGCTCAAGCGGTAAGATATGCC---CCCGTGTGTTCTGTGATTAATGCC 183
Db      157 CTCGATTTGGTGCACAAAGTCACCGAGTATAACAGTGTGTTTACCTGCTGTAGTGGC 216
QY      184 CAAAGCTCGGCTCAGTACATTCCTGATGCTGCTAAATTTCTACAAAGTGGAGCAATTCAGG 243
Db      217 CAAATATCTTCTGATTAATTTCCAGACTCGAAATTTTCAAGGTGGAAGCAATTTGTGAG 276
QY      244 CCCTGGCGAGTCTCGCAAGTTTCTCGGCTTGTCTAAATTTGTTTAAATTTGTTTCTGAGGTGTTACT 303
Db      277 CCATGAGGATCCAGCAAGTTTTCATCGCTTTACTGAAATTCGGATTCGAGGTGTTACT 336
QY      304 GTTCTGATGTTTCGAGGTTTGTGCTCAAGGTGTTCTCACTGAGAGGCAAGCGGCTCA 363
Db      337 GTTCTGATGTGAGAGGTTTGTGTCACAAAGGAGTTCTACCGAGAGACACGGTGGCTCT 396
QY      364 GAATTTTCTGAACACAGTTTGTGCTAAAGTTAAAGTGAAGATGAGATCGTGTGTTAGCAAGAC 423
Db      397 GAGTCTCGGAAGACAAATTTGTGCTAAAGTTAAAGTGAAGATCGTGTGTTAGCAAGAC 456
QY      424 CAGGTTGAGGATGTTATAGAAAAATCATTTGAGGAGGCAAGAACTGAGAGATTTGAGAGAC 483
Db      457 CAAGTGAATCTGTAATCAACACATAATTTGAAGGAGCAAGGACAGAGAGATTTGTTGAT 516
QY      484 GCGAAGATTTTCTGCTGCTGTTTCAGATGTAATAGAGTCCGCACTGGTGAAGCGGGT 543
Db      517 GCGAAGATTTTGTGCTGCTGTTTCAGATGTCATAGAGTTAGGAGTTAGCAGGTGAGCGTGG 576
QY      544 GATAAGCTGAGAGGATGACAGGAG 568
Db      577 GAGAAAGCAGAGAGATGACTGTG 601

RESULT 10
AY442185 947 bp mRNA linear PLN 15-NOV-2003
Lycopersicon esculentum PII-like protein (GLB1) mRNA, complete cds.
DEFINITION
ACCESSION AY442185
VERSION AY442185.1 GI:38231569
KEYWORDS
SOURCE
ORGANISM Lycopersicon esculentum (tomato)
Lycopersicon esculentum
REFERENCE
1 (bases 1 to 947)
AUTHORS Walch-Liu,P., Roemheld,V. and von Wiren,N.
A PII-like protein in tomato
Unpublished
JOURNAL
REFERENCE 2 (bases 1 to 947)
AUTHORS Walch-Liu,P., Roemheld,V. and von Wiren,N.
Direct Submission
JOURNAL Submitted (17-OCT-2003) Institute of Plant Nutrition, Hohenheim
University, Fruwirthstr.20, Stuttgart 70593, Germany
FEATURES
Source
1..947
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
1..947
/genes="GLB1"
/notes="LeGLB1"
53..652
CDS

```

```

/gene="GLB1"
/note="similar to nitrogen regulatory protein PII"
/codon_start=1
/product="PII-like protein"
/protein_id="AAR14689.1"
/db_xref="GI:39231570"
/translation="WASPSLSKSNFSLHSFSSPSLSQPHPTISIVVOPKPPSPQLT
VRQCNQSPFPIIRQNSPDFVDAKFYKVEAILRPWIOQVSSALLKMGIRGVTSVD
FRCGQAGLTERQAGSESDTFVAKVMEIVSXQVQGVIAKIEBARTGEIGDG
KIFLTPISDVIRVTRGEGEKAERMWGGHADMSALSTS"

```

ORIGIN

```

Query Match      46.3%; Score 275.2; DB 8; Length 947;
Best Local Similarity 75.9%; Pred. No. 2.5e-65;
Matches 340; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 135 CGCGGTCAAGCGGTGAAGATATGCCCGTCGCTCTGTGATTAATGCCCMAAGCTCGCC 194
Db 181 CACCTTCAACCGCTGCAAAATGCTCTCTTCCGATATCAGAGCCCAAACTCTCC 240

QY 195 TGACTACATTCCTGANGCTAAATCTTCAAAAGTGAAGCAATCTCAGGCCCTGGCGAGT 254
Db 241 AGATTTCGTCCTGTGTCGGAAGTTTACAAAGTTGAAGCAATCTTAGACCTTTGGAGAAT 300

QY 255 CTCGCAAGTTTCTCGCTTCTGCTAAATGTTGATTCGAGGTCTACTGTTTCTGATGT 314
Db 301 TCAACAGAGTTTCTCGCTTCTGCTAAATGTTGATTCGAGGTCTACTGTTTCTGATGT 360

QY 315 TCGAGTTTGTGCTCAAGTGTCTCACTGAGGCGAGGCGCTCGAATTTTCTGA 374
Db 361 TCGTGTGTTTGGCCCAAGTGGCTTGCTGCTGAGGCGAAGCTGCTGAAATCTCTGA 420

QY 375 AGCAAGTTTGTGCTAAAGTTAAGATGAGATCGTGGTTAGCAAGACAGAGTTGAGGA 434
Db 421 AGACAGTTTGTGCTAAAGTTAAGATGAGATCGTGGTTAGCAAGACAGAGTTGAGGA 480

QY 435 TGTATAGAAAAATCATTTGAGGAGCAAGACTGAGAGATTTGAGAGCGGCAAGATTTT 494
Db 481 AGTCATTGCAAGATAAATTTGAAGAGCAAGACTGAGAGATTTGAGAGCGGCAAGATTT 540

QY 495 CTTCGCTGCTGTTTTCAGATGTAATAGATGTCGCTGAGCGGGGTGATAGGCTGA 554
Db 541 CTTCGCTGCTGTTTTCAGATGTAATAGATGTCGCTGAGCGGGGTGATAGGCTGA 600

QY 555 GAGGATGACAGGAGGCGGATCTGACATG 582
Db 601 GAGGATGACAGGAGGCGGATCTGATATG 628

```

RESULT 11

```

AK068407
LOCUS      Oryza sativa (japonica cultivar-group) cDNA clone:J013149808, full
DEFINITION insert sequence.
ACCESSION AK068407
VERSION    AK068407.1 GI:32978425
KEYWORDS   FLI cDNA; CAP trapper.
SOURCE     Oryza sativa (japonica cultivar-group)
ORGANISM   Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzaceae; Oryza.
            1

```

REFERENCE

```

AUTHORS    The Rice Full-length cDNA Consortium, National Institute of
            Agrobiological Sciences Rice Full-length cDNA Project Team:
            Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
            Kishimoto, N., Yasaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
            Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
            Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
            Science Genome Sequencing & Analysis Group: Ootomo, Y., Murakami, K.,
            Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
            Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
            Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,

```

```

Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., Riken, J.,
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
Hara, A., Hashidume, W., Hayashi, N., Imotani, K., Ishii, Y., Itoh, M.,
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Yoshino, M. and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
Science 301 (5631), 376-379 (2003)

```

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

REFERENCES

ORIGIN	/db_xref="taxon:39947" /clone="J013149808"			
Query Match	36.4%	Score 216.2	DB 8	Length 902
Best Local Similarity	70.9%	Pred. NO. 6.1e-49		
Matches 287	Conservative 0	Mismatches 118	Indels 0	Gaps 0
QY	189	CTCGCTGACTACATTCCTGATGCTAAATTTCTACAAAGTGAAGCAATTCCTGAGCCCTG	248	
DB	267	CGCAGCAGGGTACCAGCGGAGTCGGAGTTCTTCAAGGTGGAGGCAATTCCTGAGGCCATG	326	
QY	249	GCGAGTCTCGCAAGTTTCTCTCGGCTTTCCTAAATTTGGTATTTCGAGGTGTACTGTTC	308	
DB	327	GAGGGTGCTTATGTGTCATCGGTTTCTGCAATGGGATCAGAGCGGTGAGCGGTGTC	386	
QY	309	TGATGTTTCGAGGTTTGGTGCTCAAGGTGGTTTCAACTGAGAGCGAGCGGCTCAGAAATT	368	
DB	387	CGAGTCTCGGAGTTTTCGCGACAGGCGGGTCAACTGAGAGCGCATGAAGGTCAGAAATT	446	
QY	369	TTCGAGACAGAAAGTTTGTGCTTAAAGTTAAGATGAGATCGGTGTAGCAAGACCGGT	428	
DB	447	TGCAGAGATACATTATTCGATAAGTTAAGATGGAATAGTGTGTCTCAAGGATCAGGT	506	
QY	429	TGAGATGTTATAGAAAATCATCTGAGGCGGAGCTGCGAGGATTCGAGAGCGCA	488	
DB	507	TGAAGCTGTTGTTGACAGATTAATTTGAAAGCGCAAGACAGAGAAATTTGTTGATGGAA	566	
QY	489	GATTTCTGCTGCTGTTTTCAGATGTAATAAGAGTCGCACTGTTGAGCGGGGTGATAA	548	
DB	567	AATATTTTGTATCCGCTGTCGAGCTGATCAGATACGACCGCGGAGCGGAGCG	626	
QY	549	GGCTGAGAGGATGACAGAGCGGCGATTCGATGATGATGATCTCTGC	593	
DB	627	AGCGGAGAGGATGCGCGGCGGCTGGCGGACAAAGCTGTCTCAGC	671	
RESULT 12	AK099152 902 bp mRNA linear PIN 24-JUL-2003			
LOCUS	Oryza sativa (japonica cultivar-group) cDNA clone:J023070H02, full			
DEFINITION	insert sequence.			
ACCESSION	AK099152			
VERSION	AK099152.1 GI:32984361			
KEYWORDS	F11 DNA; CAP trapper.			
SOURCE	Oryza sativa (japonica cultivar-group)			
ORGANISM	Oryza sativa (japonica cultivar-group)			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoidae; Oryzaceae; Oryza.			
AUTHORS	The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yasaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Nishiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Kojima, K., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Iida, Y., Tsunoda, Y., Narikawa, R., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Kurokawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yamada, K., Riken, Kusumegi, T., Oka, M., Ryu, Ueda, M., Matsubara, K., Riken, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Oka, Y., Saitoh, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.			
TITLE	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice			
JOURNAL	Science 301 (5631), 376-379 (2003)			
MEDLINE	22752273			
PUBMED	12869764			
REFERENCE	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hayashizaki, Y., Hara, A., Hashizume, W., Hiraoka, T., Hori, F., Hotta, I., Iida, Y., Ikeda, R., Ishikawa, M., Itoh, M., Kagawa, I., Kishikawa-Hirozane, T., Kishimoto, N., Kojima, K., Kondo, S., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nishiki, T., Ohneda, E., Ooka, M., Ooka, H., Otsu, N., Ota, Y., Saitoh, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sugabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A., and Hayashizaki, Y.			
FEATURES	Location/Qualifiers			
source	1..902			
	organism="Oryza sativa (japonica cultivar-group)"			
	mol_type="mRNA"			
	cultivar="Nipponbare"			
	db_xref="taxon:39947"			
	clone="J023070H02"			
ORIGIN	Query Match 36.4% Score 216.2 DB 8 Length 902			
	Best Local Similarity 70.9% Pred. NO. 6.1e-49			
	Matches 287 Conservative 0 Mismatches 118 Indels 0 Gaps 0			
QY	189	CTCGCTGACTACATTCCTGATGCTAAATTTCTACAAAGTGAAGCAATTCCTGAGCCCTG	248	
DB	267	CGCAGCAGGGTACCAGCGGAGTCGGAGTTCTTCAAGGTGGAGGCAATTCCTGAGGCCATG	326	

QY 249 GCGAGTCGCGAAGTTTCTCTCGGCTTTGCTTAAATAATGTTATTCGAGGTTGTTACTGTTTC 308
 DB |||||
 DB 327 GAGGTCGCTTATGTCTCATCGGTTTGTGCAAAATGCGGATCAGAGGCGTGTCGCGTGC 386
 QY 309 TGATGTCGAGGTTTGTGTCCTCAAGTGGTTCAACTGAGAGGCGAGGCGGCTCGAATTT 368
 DB |||||
 DB 387 CGAGTCGCGGTTTTCGCGGCAACGCGCGGCTCACTGAGAGGCAATGAGGTTCAAGTT 446
 QY 369 TTCTGAAGCAAGTTTGTGCTTAAAGTTAAGATGAGATCGTGGTTAGCAAAAGACAGGT 428
 DB |||||
 DB 447 TCCAGAGATACATTTATTTGATAAAGTTAAGATGGAATAGTGGTGTCCAAAGATCAGGT 506
 QY 429 TCAGAGTTATAGAAAAAATCATTTGAGAGGCGAGAACTCGAGAGATTTGGAGACGGGAA 488
 DB |||||
 DB 507 TGAAGCTGTTGTGCAAGATAAATGCAAAAGGCAAGAACAGGAGAAATTTGGTATGGAAA 566
 QY 489 GATTTTCTGCTGCTGCTTTCAGTGTATTAAGATGCGCACTGTTGAGCGGCGGTGATAA 548
 DB |||||
 DB 567 AATATTTTGTGATACCGCTGTGCGAGCTGATCAGATPACCGACGCGCAAGAGGCGG 626
 QY 549 GGCTGAGAGGATGACAGGAGGCGGATCTGACATGAGTACTTCTGC 593
 DB |||||
 DB 627 AGCGAGAGGATGCGCGAGGCGCTGCGGCAAGCTGCTCTCAGC 671

RESULT 13
 LOCUS PPI489604 1085 bp mRNA linear PLN 11-OCT-2003
 DEFINITION Pinus pinaster mRNA for PII-like protein (gln gene).
 ACCESSION AJ489604
 VERSION AJ489604.1 GI:37653226
 KEYWORDS gln gene; PII-like protein.
 SOURCE Pinus pinaster
 ORGANISM Pinus pinaster
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Coniferales; Pinaceae; Pinus; Pinus.
 1
 Cantan, F.J.
 A PII-like protein from xylem of adult Maritime pine trees
 Unpublished
 JOURNAL
 REFERENCE 2 (bases 1 to 1085)
 Cantan, F.J.
 AUTHORS
 TITLE Direct Submission
 JOURNAL Submitted (11-JUN-2002) Cantan F.J., Molecular Biology and
 Biochemistry, Faculty of Sciences, Campus de Teatinos s/n, E-29071,
 SPAIN

FEATURES
 source
 1..1085
 /organism="Pinus pinaster"
 /mol_type="mRNA"
 /db_xref="taxon:71647"
 /tissue_type="xylem"
 /dev_stage="adult tree"
 gene
 1..1085
 /gene="gln"
 CDS
 42..773
 /gene="gln"
 /function="putative nitrogen sensor protein"
 /codon_start=1
 /product="PII-like protein"
 /protein_id="CAD33967.1"
 /db_xref="GI:37653227"
 /translation="MAAHLPLFVKGSIPLSPSTSTSSCTYSISHTAASPFGHSP
 AAANKSVLGMTGVRSLKRMKQOIKAQAQAPNGTTHPDYVPEA
 NPYKVEILRPWIRISHTVTLGLKMGIRGVTVDVRGFGVQAGSAERQAGSEFSDNFV
 SKYMEIVVSKQDVEAVIDAIDDEARTGIGDKLFVFPVADVIRVTRGERDLKPERM
 AGGQRYLQVYIRLQTLATSLAL"

ORIGIN

Query Match 36.1%; Score 215.4; DB 8; Length 1085;
 Best Local Similarity 73.2%; Pred. No. 1e-48;
 Matches 276; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 193 CCTGACTACATTCCTGTCGTTCTAAATTTCTACAGTGGAGCAATTTCTCAGGCGCTCGCGA 252
 DB |||||
 DB 327 CCAGATTATGTCCAGAGCCCACTTTTACAAAGTAGAAGCAATATTGAGGCGCATGGCGC 386
 QY 253 GTCTCCAAAGTTTCTCGGCTTTCGCTTAAATAATGTTATTCGAGGTGTTACTGTTTCTGAT 312
 DB |||||
 DB 387 ATCTCCCATGCACTACGAGGTCTATTGAATAATGGGATTCGTGGCGTAATCTCTCTGAT 446
 QY 313 GTTCGAGGTTTGTGGTCTCAAGTGGTTCAACTGAGAGGCGGCGGCTCAGAAATTTTCT 372
 DB |||||
 DB 447 GTTAGAGTTTGGAGTTCAGGCTGATCTGCAGAACGGCAAGCAGGCTCCGAGTTTCT 506
 QY 373 GAAGACAAGTTTGTCTTAAAGTTAAGATGAGATGCTGTTAGCAAGACCAAGGTTGAG 432
 DB |||||
 DB 507 AAAGACAATTTTGTGTCAAAAGTAAAGATGGAGATTGTGGTATCTTAAAGATCAGGTAGAA 566
 QY 433 GATGTTATAGAAAATAATCATTTGAGGAGGCAAGAACTGGAGAGATTGGAGACGCAAGATT 492
 DB |||||
 DB 567 GAAGTATTTGATGCATCATTTGATGAGGCAAGAACTGGAGAAATTTGAGATGCAATAA 626
 QY 493 TTCTTGCTGCTGTTTTCAGATGTAATPAGAGTCCGCACTGTTGAGCGGCGGTGATAAGCT 552
 DB |||||
 DB 627 TTGTGCTTCCAGTTCCAGATGTCATTCTGTTGAGAACAGGTGAGCGGACTTGAAGCCA 686
 QY 553 GAGAGGATGACAGGAGG 569
 DB |||||
 DB 687 GAGAGAATGGCTGCTGG 703

RESULT 14
 LOCUS AF017419 686 bp DNA linear BCT 15-JUL-1998
 DEFINITION Nostoc punctiforme PII (glnB) gene, complete cds.
 ACCESSION AF017419
 VERSION AF017419.1 GI:2394287
 KEYWORDS Nostoc punctiforme
 SOURCE Nostoc punctiforme
 ORGANISM Nostoc punctiforme
 Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 REFERENCE 1 (bases 1 to 686)
 Hanson, T.E., Forchhammer, K., Tandeau de Marsac, N. and Meeks, J.C.
 AUTHORS Characterization of the glnB gene product of Nostoc punctiforme
 TITLE strain ATCC 29133: glnB or the PII protein may be essential
 JOURNAL Microbiology 144 (Pt 6), 1537-1547 (1998)
 MEDLINE 98304077
 PUBMED 9639924
 REFERENCE 2 (bases 1 to 686)
 Hanson, T.E. and Meeks, J.C.
 AUTHORS Direct Submission
 TITLE Submitted (05-AUG-1997) Section of Microbiology, University of
 JOURNAL California, One Shields Avenue, Davis, CA 95616, USA
 FEATURES
 Location/Qualifiers
 1..686
 /organism="Nostoc punctiforme"
 /mol_type="genomic DNA"
 /strain="ATCC29133; PCC 73102"
 /db_xref="ATCC:29133"
 /db_xref="taxon:63737"
 275..613
 /gene="glnB"
 275..613
 /gene="glnB"
 /function="putative nitrogen regulatory protein"
 /codon_start=1
 /transl_table=11
 /product="PII"
 /protein_id="AAC26348.1"
 /db_xref="GI:2394288"

ORIGIN

QY 193 CCTGACTACATTCCTGTCGTTCTAAATTTCTACAGTGGAGCAATTTCTCAGGCGCTCGCGA 252
 DB |||||
 DB 327 CCAGATTATGTCCAGAGCCCACTTTTACAAAGTAGAAGCAATATTGAGGCGCATGGCGC 386
 QY 253 GTCTCCAAAGTTTCTCGGCTTTCGCTTAAATAATGTTATTCGAGGTGTTACTGTTTCTGAT 312
 DB |||||
 DB 387 ATCTCCCATGCACTACGAGGTCTATTGAATAATGGGATTCGTGGCGTAATCTCTCTGAT 446
 QY 313 GTTCGAGGTTTGTGGTCTCAAGTGGTTCAACTGAGAGGCGGCGGCTCAGAAATTTTCT 372
 DB |||||
 DB 447 GTTAGAGTTTGGAGTTCAGGCTGATCTGCAGAACGGCAAGCAGGCTCCGAGTTTCT 506
 QY 373 GAAGACAAGTTTGTCTTAAAGTTAAGATGAGATGCTGTTAGCAAGACCAAGGTTGAG 432
 DB |||||
 DB 507 AAAGACAATTTTGTGTCAAAAGTAAAGATGGAGATTGTGGTATCTTAAAGATCAGGTAGAA 566
 QY 433 GATGTTATAGAAAATAATCATTTGAGGAGGCAAGAACTGGAGAGATTGGAGACGCAAGATT 492
 DB |||||
 DB 567 GAAGTATTTGATGCATCATTTGATGAGGCAAGAACTGGAGAAATTTGAGATGCAATAA 626
 QY 493 TTCTTGCTGCTGTTTTCAGATGTAATPAGAGTCCGCACTGTTGAGCGGCGGTGATAAGCT 552
 DB |||||
 DB 627 TTGTGCTTCCAGTTCCAGATGTCATTCTGTTGAGAACAGGTGAGCGGACTTGAAGCCA 686
 QY 553 GAGAGGATGACAGGAGG 569
 DB |||||
 DB 687 GAGAGAATGGCTGCTGG 703

Query Match 18.0%; Score 107; DB 1; Length 686;
 Best Local Similarity 60.4%; Pred. No. 1.3e-18;
 Matches 195; Conservative 0; Mismatches 125; Indels 3; Gaps 1;

QY 214 AAATTCACAAAGTGGAGCAATTCCTCAGGCCCTCGGAGTCTCGCAAGTTTCCTCGGCT 273
 DB 272 AAGATGAAAGAGTAGAGCTATTATCCGCCCATTTAAGCTAGATGAGTAAAGATTGCC 331

QY 274 TTCTCTAAATTTGGTATTCGAGGTGTTACTGTTTCTGATGTTTCAGGTTTGGTGTCAA 333
 DB 332 TTGGTTAACTGCTGTTATGTCGATGCTGTTCTGAAAGTTTCGGGGTTTCGACGCA 391

QY 334 GTTGTTCACACTGAGAGCGAGGCGGCTCAGATTTTCTGAAGACAAAGTTTGTGCTAAA 393
 DB 392 AAGAGGCCAAACTGAACGGTATCGGGTTCAGTACACCGTTG---AGTTTTCGAAAA 448

QY 394 GTTAAGATGAGATCGTGGTTAGCAAGACCCAGGTTGAGGATGTTATAGAAAAATCATT 453
 DB 449 CTCNAAGTGAATCGTCTGTGACCAATCAGGTTGATATGGTGTAGACAAAATTATT 508

QY 454 GAGGAGCAAGAACTGAGAGATTGGAGAGCGCAAGATTTTCTGCTGCTGCTGTTTCAGT 513
 DB 509 GCTGCTCCCGCACTGTTGAAATCGGCGATGGTAAATTTTCACTCGCCTGTTGAGCAA 568

QY 514 GTTAATAGAGTCCGACTGGTGA 536
 DB 569 GTGATTCGATTCGATTCGCGCA 591

RESULT 15
 ASP251822 2493 bp DNA linear BCT 15-FEB-2000
 LOCUS Anabaena sp. PCC7120 glnB gene for PII protein.
 DEFINITION Anabaena sp. PCC7120 glnB gene for PII protein.
 ACCESSION AJ251822
 VERSION AJ251822.1 GI:6983586
 KEYWORDS glnB gene; PII protein.
 SOURCE Nostoc sp. PCC 7120 (Anabaena sp. PCC 7120)
 ORGANISM Nostoc sp. PCC 7120
 Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 REFERENCE 1
 Gonzalez, L., Phalip, V. and Zhang, C.C.
 Phosphorylation of the signal transduction protein PII by the
 Ser/Thr kinase PknC in the cyanobacterium Anabaena sp. strain PCC
 7120
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2493)
 AUTHORS Zhang, C.C.
 TITLE Direct Submission
 JOURNAL Submitted (14-DEC-1999) Zhang C.C., Laboratoire de Chimie
 Bacterienne, C.N.R.S., 31, chemin Joseph Aiguier, 13402 Marseille
 cedex 20, FRANCE
 FEATURES
 source Location/Qualifiers
 1..2493
 /organism="Nostoc sp. PCC 7120"
 /mol_type="genomic DNA"
 /strain="PCC 7120"
 /db_xref="taxon:103690"
 1074..1412
 /gene="glnB"
 1074..1412
 /gene="glnB"
 /function="signaling protein"
 /codon_start=1
 /transl_table=11
 /product="PII protein"
 /protein_id="CAB75358.1"
 /db_xref="GI:6983587"
 /db_xref="GOA:Q9L422"
 /db_xref="SPTRMBL:Q9L422"
 /translation="MKKVEAIRPFLDEVKIALNAGIVMTVSEVGRGKQKQTE
 RYRGSYTVTFELQKLKVEIVVEDNQVMDVKIIAAARTGEIDGKIFISPEVQIRI
 RTGKNTAEV"

ORIGIN

Query Match 17.4%; Score 103.2; DB 1; Length 2493;
 Best Local Similarity 58.5%; Pred. No. 1.7e-17;
 Matches 199; Conservative 0; Mismatches 138; Indels 3; Gaps 1;

QY 216 ATTCTACAAAGTGGAGCAATTCCTCAGGCCCTCGGAGTCTCGCAAGTTTCCTCGGCTTT 275
 DB 1073 AATGAAAAAGTAGAGCTATTATCCGCCCATTTAAGCTAGATGAGTAAAGATTGCCCTTT 1132

QY 276 GCTAAAAATTGGTATTCGAGGTGTTACTGTTTCTGATGTTTCGAGTTTGGTGTCAAGG 335
 DB 1133 AGTCAACCGCGGTATTGTGGGTATGACTGTTTCTGAAAGTCCGGGTTTGGACGCGCAA 1192

QY 336 TGGTTCAACTGAGAGCGAGGCGGCTCAGATTTTCTGAAGACAAAGTTTGTCTAAAAGT 395
 DB 1193 AGGACAGACAGAACGCTATCGCGGCTCTGACTACACTGTGG---AGTTTCTGAAAAACT 1249

QY 396 TAAGATGGAGATCGTGGTTAGCAAGACCCAGGTTGAGGATGTTATAGAAAAATCATTGA 455
 DB 1250 GAAGGTGGAGATTCTAGTTGAAGATAACCAAGTTGATATGGTCTTGACAAAAATCATTGC 1309

QY 456 GGAGGCAAGAACTCGAGAGATTGGAGAGCGCAAGATTTTCTGCTGCTGCTGTTTCAGATGT 515
 DB 1310 TGCTGCTCGTACAGGCGAAATCGGTGATGGTAAATCTTTATCTCACCTGTAGACAAAGT 1369

QY 516 AATAAGAGTCCGCACTGTGTGAGCGGGTGATAAGGCTGAG 555
 DB 1370 AATTCGGATTCTGACTGTGAGAGAGAAATACAGAGAGCCCGTG 1409

Search completed: May 27, 2004, 22:22:46
 Job time : 1758 secs

Result No.	Score	Query %		Length	DB	ID	Description
		Match	Length				
1	594	100.0	594	4	AAF58584	AAF58584	Ricinus c
2	594	100.0	597	4	AAF58582	AAF58582	Ricinus c
3	285	48.0	588	4	AAF58583	AAF58583	Arabidops
4	285	48.0	591	6	AB214328	AB214328	Arabidops
5	285	48.0	817	4	AAF58581	AAF58581	Arabidops
6	96.4	16.2	110000	2	AAV21209_01	Continuation (2 of	
7	96.2	16.2	110000	2	AAV21209_14	Continuation (15 of	
8	84.8	14.3	96109	4	AAF28548	AAF28548	Genomic f
9	83.6	14.1	375	8	ADA330024	ADA330024	DNA encod
10	81.4	13.7	339	7	ACF72027	ACF72027	Photorhab
11	81.4	13.7	110000	7	ACF67367_52	Continuation (53 of	
12	81.4	13.7	110000	7	ACF65387_3	Continuation (4 of	
13	77.6	13.1	110000	2	AA42063_03	Continuation (4 of	
14	77	13.0	363	7	ABZ40022	ABZ40022	N. gonorr
15	73	12.3	16526	3	AA381472	AA381472	N. mening
16	73	12.3	110000	3	AA381489_6	Continuation (7 of	
17	73	12.3	172325	3	AAF21613_1	AAF21613	Neisseria
18	73	12.3	172325	3	AAF21612	AAF21612	Neisseria
19	70.8	11.9	336	5	AAH68445	AAH68445	C glutami
20	70.8	11.9	417	5	AAH67222	AAH67222	C glutami
21	70.8	11.9	349980	5	AAH68531	AAH68531	C glutami
22	64.8	10.9	290	4	AAF71296	AAF71296	Corynebac
23	64.2	10.8	1164	5	AA588441	AA588441	DNA encod

SQ Sequence 594 BP; 149 A; 117 C; 157 G; 171 T; 0 U; 0 Other;
 Query Match 100.0%; Score 594; DB 4; Length 594;
 Best Local Similarity 100.0%; Pred. No. 2.1e-168;
 Matches 594; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGGCACGAGCTACTGCGAACTGGCTTGTCTCACTCTCTTCTTCTAATAACATCAAG 60
 DB 1 CGGCACGAGCTACTGCGAACTGGCTTGTCTCACTCTCTTCTTCTAATAACATCAAG 60
 QY 61 AAGAAATTCCTGTTTGTATTCAGTTTGTTCAGAGCTTAGACATTTCTCGTTT 120
 DB 61 AAGAAATTCCTGTTTGTATTCAGTTTGTTCAGAGCTTAGACATTTCTCGTTT 120
 QY 121 TCTCACTTTAAACACCGGGTCAAGCGCTTAAGATATGCCCGTCTTCTGTAATTAAT 180
 DB 121 TCTCACTTTAAACACCGGGTCAAGCGCTTAAGATATGCCCGTCTTCTGTAATTAAT 180
 QY 181 GCCCAAGAGCTGCTGACTACATTCCTGATGCTAAATTTCTCAAGTGGAGCAATTC 240
 DB 181 GCCCAAGAGCTGCTGACTACATTCCTGATGCTAAATTTCTCAAGTGGAGCAATTC 240
 QY 241 AGGCCCTGGCGAGTCTCGCAAGTTTCTCGGCTTGTCTAATAATTTGTAATTCGAGTGT 300
 DB 241 AGGCCCTGGCGAGTCTCGCAAGTTTCTCGGCTTGTCTAATAATTTGTAATTCGAGTGT 300
 QY 301 ACTGTTCTGATGTTTCGAGGTTTGGTGTCTCAAGTGGTTCAACTGAGAGGCGGCGG 360
 DB 301 ACTGTTCTGATGTTTCGAGGTTTGGTGTCTCAAGTGGTTCAACTGAGAGGCGGCGG 360
 QY 361 TCAGAAATTTCTGAAGCAAGTTTGTCTAAGTAAAGTAAAGTAAAGTAAAGTAAAG 420
 DB 361 TCAGAAATTTCTGAAGCAAGTTTGTCTAAGTAAAGTAAAGTAAAGTAAAGTAAAG 420
 QY 421 GACCAAGTTGAGGATGTTATAGAAAATCAATTCAGGAGGCAAGAACTGGAGATTGGA 480
 DB 421 GACCAAGTTGAGGATGTTATAGAAAATCAATTCAGGAGGCAAGAACTGGAGATTGGA 480
 QY 481 GACGGAAGATTTTCTGCTGCTTTCAGATGTAATAAGTCCGCACTGGTGAGCGG 540
 DB 481 GACGGAAGATTTTCTGCTGCTTTCAGATGTAATAAGTCCGCACTGGTGAGCGG 540
 QY 541 GGTGATAGGCTGAGAGGATGACAGGAGGCGGATCTGACATGAGTACTTCTGCT 594
 DB 541 GGTGATAGGCTGAGAGGATGACAGGAGGCGGATCTGACATGAGTACTTCTGCT 594

RESULT 2
 AAF58582
 ID AAF58582 standard; cDNA; 897 BP.
 AC AAF58582;
 XX
 DT 23-APR-2001 (first entry)
 XX
 DE Ricinus communis P-II cDNA.
 XX
 KW Castor bean; PII; plant nitrogen regulatory gene; P-II;
 FW nitrogen assimilation; transgenic plant; herbicide screening; es.
 XX
 OS Ricinus communis.
 XX
 EN US6177275-B1.
 XX
 PD 23-JAN-2001.
 XX
 PF 23-JUL-1997; 97US-00899330.
 XX
 PR 24-JUL-1996; 96US-0022328P.
 XX
 PA (UNY) UNIV NEW YORK STATE.
 XX
 PI Coruzzi GM, Lam H, Hsieh M;

XX WPI; 2001-158572/16.
 DR P-FSDS; AAB69496.
 XX
 PT Novel P-II genes capable of regulating plant nitrogen assimilation,
 PT useful for transgenic plant production, and as probes for isolating
 PT additional genomic clones having P-II gene promoters.
 XX
 PS Claim 2; Fig 13; 35pp; English.
 XX
 CC The present sequence encodes a nitrogen regulatory PII protein. Novel
 CC plant PII (also called P-II) nucleotide sequences have been isolated.
 CC They are useful for regulating nitrogen assimilation in plants, and in
 CC transgenic plant production. They are also used to engineer organisms
 CC that overexpress wild-type or mutant P-II regulatory proteins. P-II
 CC proteins are useful for in vitro screening of herbicides. P-II
 CC nucleotides may be used as probes for isolating additional genomic clones
 CC with the promoters of P-II genes. P-II promoters are light- and/or
 CC sucrose-inducible, and are suitable for genetic engineering of plants
 XX
 SQ Sequence 897 BP; 254 A; 171 C; 218 G; 254 T; 0 U; 0 Other;
 Query Match 100.0%; Score 594; DB 4; Length 897;
 Best Local Similarity 100.0%; Pred. No. 2.5e-168;
 Matches 594; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGGCACGAGCTACTGCGAACTGGCTTGTCTCACTCTCTTCTTCTAATAACATCAAG 60
 DB 50 CGGCACGAGCTACTGCGAACTGGCTTGTCTCACTCTCTTCTTCTAATAACATCAAG 109
 QY 61 AAGAAATTCCTGTTTGTATTCAGTTTGTTCAGAGCTTAGACATTTCTCGTTT 120
 DB 110 AAGAAATTCCTGTTTGTATTCAGTTTGTTCAGAGCTTAGACATTTCTCGTTT 169
 QY 121 TCTCACTTTAAACACCGGGTCAAGCGCTTAAGATATGCCCGTCTTCTGTAATTAAT 180
 DB 170 TCTCACTTTAAACACCGGGTCAAGCGCTTAAGATATGCCCGTCTTCTGTAATTAAT 229
 QY 181 GCCCAAGAGCTGCTGACTACATTCCTGATGCTAAATTTCTCAAGTGGAGCAATTC 240
 DB 230 GCCCAAGAGCTGCTGACTACATTCCTGATGCTAAATTTCTCAAGTGGAGCAATTC 289
 QY 241 AGGCCCTGGCGAGTCTCGCAAGTTTCTCGGCTTGTCTAATAATTTGTAATTCGAGTGT 300
 DB 290 AGGCCCTGGCGAGTCTCGCAAGTTTCTCGGCTTGTCTAATAATTTGTAATTCGAGTGT 349
 QY 301 ACTGTTCTGATGTTTCGAGGTTTGGTGTCTCAAGTGGTTCAACTGAGAGGCGGCGG 360
 DB 350 ACTGTTCTGATGTTTCGAGGTTTGGTGTCTCAAGTGGTTCAACTGAGAGGCGGCGG 409
 QY 361 TCAGAAATTTCTGAAGCAAGTTTGTCTAAGTAAAGTAAAGTAAAGTAAAGTAAAG 420
 DB 410 TCAGAAATTTCTGAAGCAAGTTTGTCTAAGTAAAGTAAAGTAAAGTAAAGTAAAG 469
 QY 421 GACCAAGTTGAGGATGTTATAGAAAATCAATTCAGGAGGCAAGAACTGGAGATTGGA 480
 DB 470 GACCAAGTTGAGGATGTTATAGAAAATCAATTCAGGAGGCAAGAACTGGAGATTGGA 529
 QY 481 GACGGAAGATTTTCTGCTGCTTTCAGATGTAATAAGTCCGCACTGGTGAGCGG 540
 DB 530 GACGGAAGATTTTCTGCTGCTTTCAGATGTAATAAGTCCGCACTGGTGAGCGG 589
 QY 541 GGTGATAGGCTGAGAGGATGACAGGAGGCGGATCTGACATGAGTACTTCTGCT 594
 DB 590 GGTGATAGGCTGAGAGGATGACAGGAGGCGGATCTGACATGAGTACTTCTGCT 643
 RESULT 3
 AAF58583
 ID AAF58583 standard; cDNA; 588 BP.
 XX
 AC AAF58583;
 XX

```

DT 23-APR-2001 (first entry)
XX Arabidopsis thaliana P-PII cDNA fragment.
DE Arabidopsis thaliana P-PII cDNA fragment.
XX Arabidopsis thaliana P-PII cDNA fragment.
KW Arabidopsis thaliana; PII; plant nitrogen regulatory gene; P-PII;
KW nitrogen assimilation; transgenic plant; herbicide screening; ss.
OS Arabidopsis thaliana.
XX Arabidopsis thaliana.
XX US6177275-B1.
XX 23-JAN-2001.
XX 23-JUL-1997; 97US-00899330.
XX 24-JUL-1996; 96US-0022328P.
XX (UWNY ) UNIV NEW YORK STATE.
XX Coruzzi GM, Lam H, Hsieh Y;
XX WPI; 2001-158572/16.
XX P-PSDB; AAB69495.
XX Novel P-PII genes capable of regulating plant nitrogen assimilation,
XX useful for transgenic plant production, and as probes for isolating
XX additional genomic clones having P-PII gene promoters.
XX Claim 2; Col 37-38; 35pp; English.
XX The present sequence encodes a nitrogen regulatory PII protein. Novel
XX plant PII (also called P-PII) nucleotide sequences have been isolated.
XX They are useful for regulating nitrogen assimilation in plants, and in
XX transgenic plant production. They are also used to engineer organisms
XX that overexpress wild-type or mutant P-PII regulatory proteins. P-PII
XX proteins are useful for in vitro screening of herbicides. P-PII
XX nucleotides may be used as probes for isolating additional genomic clones
XX with the promoters of P-PII genes. P-PII promoters are light- and/or
XX sucrose-inducible, and are suitable for genetic engineering of plants
XX
XX Sequence 588 BP; 164 A; 105 C; 147 G; 172 T; 0 U; 0 Other;
XX
Query Match 48.0%; Score 285; DB 4; Length 588;
Best Local Similarity 70.3%; Pred. No. 2.5e-75;
Matches 397; Conservative 0; Mismatches 165; Indels 3; Gaps 1;

QY 7 GAGGCTACTGCGAACTGGGCTTGCTCCTCCTTCATCTTAATCAATCAAGAAAGAA 66
Db 7 GCGTCAATGACGAAACCCATCTCAATAACTCTCTCGGTTTCTATTCTGATCGAAGAAC 66

QY 67 TTCCCTGTTTTCGATTTGCTGTTTTCGTCAGAGCTTAGACATCTCGGTTTCTCAC 126
Db 67 ATTGCTTCTCTGATGTCATTTGATTTGCTGGAATTCAGACATTCGCCACCATTTGC 126

QY 127 TTAAACACCGCGGTCAAGCGCGTAAGATATGCC---CCGCTCGTCTCTGTGATTAATGCC 183
Db 127 CTCGATTTGTCACAAAGTCACCGAGTAATACAGTCGTTTACTCTGCTAGTGCC 186

QY 184 CAAAGTCGCGTACATCTCTGATGCTAAATTTCTACAAAGTGAAGCAATCTCAGG 243
Db 187 CAAATATCTCTGATTTATATTCAGACTCGAAATTTTACAAAGTGAAGCAATTTGCAGA 246

QY 244 CCTGGGAGTCTCGCAAGTTTCTCGGCTTTCCTAAATTTGATTCGAGGTGTACT 303
Db 247 CCATGAGAAATCAGCAAGTTTCTATCGGCTTTACTGAAATTCGGATTCGAGGTGTACT 306

QY 304 GTTCTCTGATTTGAGGTTTGTGCTCAAGTGTGTTCAACTCAGAGCGAGGCGGCTCA 363
Db 307 GTTCTCTGATGAGAGGTTTGTGCAAGAGGTTTCTACCGAGAGACACGCTGCTCT 366

QY 364 GAATTTCTGAAGACAAATTTGTTGCTAAAGTTTAAGATGAGATCTGTTAGCAAGAC 423
Db 367 GAGTTCTCGAAGACAAATTTGTTGCTAAAGTTTAAGATGAGATCTGTTAGCAAGAC 426

```

```

QY 424 CAGGTTGAGGATGTTATAGAAAATCATTGAGAGCGCAGAACTGGAGAGATTGGAGAC 483
Db 427 CAACTGGAATCTGTAAATCAACATATTTGAGAGGCGCAGGACAGGAGATTGGTAT 486

QY 484 GGCAGAGATTTTCTGCTGCTGCTGTTTCAGATGTAATAGAGTCCGCACTGGTGAGCGGGT 543
Db 487 GGCAGAGATTTTCTGCTGCTGCTGTTTCAGATGTAATAGAGTCCGCACTGGTGAGCGGGT 546

QY 544 GATAGGCTGAGAGGATGACAGGAG 568
Db 547 GAGAAAGCAGAGAGATGACTGGTG 571

RESULT 4
ABZ14328
ID ABZ14328 standard; DNA; 591 BP.
XX
XX ABZ14328;
XX
XX 21-JAN-2003 (first entry)
XX
XX Arabidopsis thaliana stress regulated gene SEQ ID NO 2133.
XX
XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX Arabidopsis thaliana.
XX
XX WO200216655-A2.
XX
XX 28-FEB-2002.
XX
XX 24-AUG-2001; 2001WO-US026685.
XX
XX 24-AUG-2000; 2000US-0227866P.
XX
XX 26-JAN-2001; 2001US-0264647P.
XX
XX 22-JUN-2001; 2001US-0300111P.
XX
XX (SCRI ) SCRIPPS RES INST.
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
XX Harper JF, Kreps J, Wang X, Zhu T;
XX
XX WPI; 2002-304127/34.
XX
XX Identifying a stress condition to which a plant cell has been exposed and
XX producing plants with increased tolerance to these abiotic stresses.
XX
XX Claim 144; SEQ ID NO 2133; 577pp + Sequence Listing; English.
XX
XX The invention relates to identifying a stress condition to which a plant
XX cell has been exposed, comprising: (a) contacting nucleic acid
XX representative of expressed polynucleotides in the plant cell with an
XX array or probes representative of the plant cell genome; and (b)
XX detecting a profile of expressed polynucleotides in the plant cell
XX characteristic of a stress response. The method is useful in the
XX production of transgenic plants, cells and seeds and in producing plants
XX with increased tolerance to abiotic stress. The present sequence is that
XX of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
XX in methods of the invention. Note: The sequence data for this patent is
XX not represented in the printed specification but is based on sequence
XX information supplied to Derwent by the European Patent Office
XX
XX Sequence 591 BP; 165 A; 105 C; 148 G; 173 T; 0 U; 0 Other;
XX
Query Match 48.0%; Score 285; DB 6; Length 591;
Best Local Similarity 70.3%; Pred. No. 2.5e-75;
Matches 397; Conservative 0; Mismatches 165; Indels 3; Gaps 1;

QY 7 GAGGCTACTGCGAACTGGGCTTGCTCCTCCTTCATCTTAATCAATCAAGAAAGAA 66
Db 7 GCGTCAATGACGAAACCCATCTCAATAACTCTCTCGGTTTCTATTCTGATCGAAGAAC 66

```


WP AAV21209_11 1100001 1210000
WP AAV21209_12 1200001 1310000
WP AAV21209_13 1300001 1410000
WP AAV21209_14 1400001 1510000
WP AAV21209_15 1500001 1610000
WP AAV21209_16 1600001 1664976

Query Match 16.2%; Score 96.4; DB 2; Length 110000;
Best Local Similarity 57.9%; Pred. No. 1.7e-17;
Matches 191; Conservative 0; Mismatches 136; Indels 3; Gaps 1;

QY 214 AATTTCTACAAAGTGGAGCAATTTCTCAGCGCTCGAGTCTCGAAGTTTCTCGGCT 273
DB 80891 AAAATGAAAAAGTTGAGCAATCATAGACCGAGAGTGGAGATTGTTAAAAAGCT 80950

QY 274 TTGCTAAAAATTTGATTCGAGGTGTACTGTTCTGATGTTTCGAGGTTTTCGTCTCAA 333
DB 80951 TTGCTGATGCTGGATATGTTGGAATGACTGTTAGTGAGGTTAAAGGTTAGGGGAGTTCAA 81010

QY 334 GGTGTTCACTGAGAGCGGCGGCTCAGAAATTTCTGAGAGCAAGTTTGTGCTAAA 393
DB 81011 GGTGGAATAGTTGAGAGGTATAGGGGAGAGATATATTGTTG---ATTAAATTCANAG 81067

QY 394 GTTAAGATGAGATCGTGGTTAGCAAGACCGAGGTGAGATGTTATAGAAAAATCAIT 453
DB 81068 GTTAAGATGATGTTGGTTGTAAGAGGAGATGTTGATATGTTATGATATATATGC 81127

QY 454 GAGGAGGCAAGAACTGGAGAGATTGGAGCGGCAAGATTTCTGCTGCTGTTTCAGAT 513
DB 81128 GAGATGCAAGAACAGCAAGCAACCGAGGATGGAATAATCTTCGTATACAGTAGAAGA 81187

QY 514 GTAATAAGATCCGACTGTTGAGCGGGT 543
DB 81188 GTCGTAAGAGTAAGAACAAAAAGAGGGT 81217

RESULT 7
AAV21209_14/c
Continuation (15 of 17) of AAV21209 from base 1400001 (Methanococcus jannaschii circular
WP Sequence split into 17 fragments LOCUS AAV21209 Accession Aav21209

WP Fragment Name Begin End
WP AAV21209_00 1 110000
WP AAV21209_01 100001 210000
WP AAV21209_02 200001 310000
WP AAV21209_03 300001 410000
WP AAV21209_04 400001 510000
WP AAV21209_05 500001 610000
WP AAV21209_06 600001 710000
WP AAV21209_07 700001 810000
WP AAV21209_08 800001 910000
WP AAV21209_09 900001 1010000
WP AAV21209_10 1000001 1110000
WP AAV21209_11 1100001 1210000
WP AAV21209_12 1200001 1310000
WP AAV21209_13 1300001 1410000
WP AAV21209_14 1400001 1510000
WP AAV21209_15 1500001 1610000
WP AAV21209_16 1600001 1664976

Query Match 16.2%; Score 96.2; DB 2; Length 110000;
Best Local Similarity 56.3%; Pred. No. 1.9e-17;
Matches 201; Conservative 0; Mismatches 153; Indels 3; Gaps 1;

QY 195 TCATACATTCCTGATGCTAAATTTCTACAAAGTGGAGCAATTTCTCAGCGCTCGGAGT 254
DB 13404 TTATAAAATGAGTTGAGATTATGAAAAAGTTGAAGCAATCATAGACCGGAGAGTT 13345

QY 255 CTCGCAAGTTTCTCGGCTTGTAAATTTGATTCGAGGTTTACTGTTCTGATCT 314
DB 13344 GGAGATTGTTAAAAAGGCTTTGTTCTGATGCTGGGTATGTTGGAATGACTGTTAGTGGT 13285

QY 315 TCGAGGTTTTGTTGCTCAAGGTGTTTCAACTGAGAGCGAGGGGCTCGAATTTTCTGA 374

DB 13284 TAAGGTAGGGAGTTTCAAGGTGGAATAGTTGAGAGGTATAGGGGAGAGATATATTGT 13225

QY 375 AGCAAGATTGTTGCTTAAAGTTTAAAGTTGAGATGAGATCGTGGTTAGCAAGACCAAGCTTGAGGA 434

DB 13224 TG---ATTTAATTTCCAAAGGTTTAAAGTTGAGTTGTTGTTTAAAGAGGAGATGTTGATAA 13168

QY 435 TGTATAGAAAAAATCATTTGAGGAGCGCAAGAACTCGGAGAGATTGGAGACGCGCAAGATTTT 494

DB 13167 TGTATTGATATCATATGCGAGATGCAAGAACAGGAAACCCAGGAGATGGAATAATCTT 13108

QY 495 CTTGCTGCTGTTTTCAGATGTTAATAGAGTCCGACTGTTGAGCGGGTCTATAAGGC 551

DB 13107 CGTCATACCAGTAGAAGAGTCTGTAAGAGTTAACAACAAGAAAGAGGAAAGAGGC 13051

RESULT 8
AAF28548
ID AAF28548 standard; DNA; 96109 BP.

XX AAF28548;

AC AAF28548;

XX 04-APR-2001 (first entry)

XX Genomic fragment #35.

XX DE DE

XX Genomic library; bacteria; human upper airway; otitis media; sinusitis;
XX bronchopulmonary; endocarditis; meningitis; ss.

XX Moraxella catarrhalis.

XX WO200078968-A2.

XX 28-DEC-2000.

XX 16-JUN-2000; 2000WO-US016649.

XX 18-JUN-1999; 99US-0140121P.

XX (INCY-) INCYTE GENOMICS INC.

XX Lagace RE, Patterson C, Berg KL;

XX WPI; 2001-041427/05.

XX Genomic library for identifying diagnostic and therapeutic compositions,
XX and for identifying virulence factors, regulatory elements and drug
XX targets, comprises Moraxella catarrhalis nucleic acids.

XX Claim 1; Page 345-368; 545pp; English.

XX The present invention relates to a Moraxella catarrhalis genomic library
XX comprising a combination of 41 nucleic acid molecules (see AAF2854-
XX AAF28554). The library has a number of uses described in the
XX specification e.g. is useful for identifying diagnostic and therapeutic
XX compositions. M. catarrhalis (Branhamella catarrhalis) is a large
XX aerobic, gram-negative diplococcus, normally found among the bacterial
XX flora of human upper airways. M. catarrhalis is known to cause acute,
XX localised infections such as otitis media, sinusitis and bronchopulmonary
XX infection and life-threatening, systemic diseases including endocarditis
XX and meningitis

XX Sequence 96109 BP; 28783 A; 18910 C; 20341 G; 28075 T; 0 U; 0 Other;

SQ

Query Match 14.3%; Score 84.8; DB 4; Length 96109;
Best Local Similarity 56.2%; Pred. No. 4.9e-14;
Matches 180; Conservative 0; Mismatches 137; Indels 3; Gaps 1;

QY 232 GCAATTCAGGCGCTCGGAGTCTCGAAGTTTCTCGGCTTTCTGCTTAAATTTGTTAT 291

DB 12953 GCAATTCAGGCGCTTTTAACTCGATGATGTCGAGAGAGCACTCTCAGAAATTCGCTC 13012

QY 292 CGAGGTGTTTACTGTTTCTGATGTTTCAGGTTTTCGCTGCTCAAGTGGTTCACACTGAGG 351

Db 13013 AATGGTATCACCGTCACTGAAGTCBAAGGCTTTGGTCGCCAAAAGGTGCATACCGAGATG 13072

Qy 352 CAGCGCGGTCTCAGAAATTTCTGAAGACAAAGTTTGTCTTAAGTTAAGATGAGAGATCGTG 411

Db 13073 TATCGTGGGCGGAATATGTGGTGTAT---TTTTCACCAAAATTAATAATTTGAGATAGCA 13129

Qy 412 GTTAGCAAGACACAGGTTGAGGATGTTATAGAAAATATCATGAGGAGGCAAGAACTGGA 471

Db 13130 TGTGCGTATGAGATGTTGATCAATTAATGAGTCAATCAATTAAGTTGCAATACAGT 13189

Qy 472 GAGATTGGAGACGCGCAAGATTTCTTCTGCTCCTGTTTCAGATGTAATAAGATCCGCACT 531

Db 13190 AATATTGCTATGTTAGATTTTGTAGTCCGCTTGAGCGTGTCAATTCGCACTTGA 13249

Qy 532 GGTGAGCGGGTGTAAAGGC 551

Db 13250 GGCATATTGATGAAGTGC 13269

RESULT 9

ADA30024

ID ADA30024 standard; DNA; 375 BP.

XX AC ADA30024;

XX AC ADA30024;

DT 20-NOV-2003 (first entry)

DE DNA encoding Acinetobacter baumannii protein #1311.

XX ds; gene; Acinetobacter baumannii; bacterial disease; antibacterial;

KW vaccine; plant biocontrol agent.

XX Acinetobacter baumannii.

XX US6562958-B1.

XX 13-MAY-2003.

XX 04-JUN-1999; 99US-00328352.

XX 09-JUN-1998; 98US-0088701P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Breton G, Bush D;

PI WPI; 2003-576092/54.

DR P-PSDB; ADA34150.

XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents

PT for diagnosing a bacterial disease, as components of antibacterial

PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for

PT plants.

XX Example; SEQ ID NO 1311; 328pp; English.

PS The invention relates to isolated Acinetobacter baumannii nucleic acids.

XX The A. baumannii nucleic acids and polypeptides are useful as reagents

CC for diagnosing a bacterial disease, as components of antibacterial

CC vaccines, as targets for antibacterial drugs, to detect the presence of

CC A. baumannii and other Acinetobacter species in a sample, in screening

CC compounds for the ability to interfere with the A. baumannii life cycle

CC or to inhibit A. baumannii infection, and as biocontrol agents for

CC plants. The present sequence represents DNA encoding an A. baumannii

CC protein.

XX Sequence 375 BP; 122 A; 64 C; 93 G; 96 T; 0 U; 0 Other;

Qy Query Match 14.1%; Score 83.6; DB 8; Length 375;

Best Local Similarity 55.9%; Pred. No. 9.5e-15;

Matches 180; Conservative 0; Mismatches 139; Indels 3; Gaps 1;

Qy 232 GCAATTTCTCAGGCCCTGGGAGTCTCGCAAGTTTCTCGGCTTTCGTAATAAATTTGTTAT 291

Db 52 GCAATTTCTCAGGCCCTGGGAGTCTCGCAAGTTTCTCGGCTTTCGTAATAAATTTGTTAT 111

Qy 292 CGAGGTGTTACTGTTTCTGATGTTTCGAGGTTTTCGTTCTCAAGTGGTTCAACTGAGAGG 351

Db 112 CAAGGGATTAACCGTAACCTGAAGTTTAAAGGTTTTCGTTCTCAAAAGGACATACAGAACTT 171

Qy 352 CAGCGCGCTCAGAAATTTCTGAAAGACAAAGTTTGTCTTAAAGTTAAGATCGAGATCGTG 411

Db 172 TACCGCGCGCTGATGATGTTGAT---TTCTTACTTAAGTAAATAATCGAAATTCG 228

Qy 412 GTTAGCAAGACACAGGTTGAGGATGTTATAGAAAATATCATGAGGAGGCAAGAACTGGA 471

Db 229 ATTAGTATGATGAATGGTTCGACGCGTAAATGAGTCAATTACACGTTGTCGCAAGCACTGGA 288

Qy 472 GAGATTGGAGACGCGCAAGATTTCTTCTGCTCCTGTTTCAGATGTAATAAGATCCGCACT 531

Db 289 AAAATCGCGACGCGTAAAGTTTGTGACTAATCTGGAACCAAGTCCGATCCGTATCCGTACA 348

Qy 532 GGTGAGCGGGTGTAAAGGC 553

Db 349 GGTGAACACAGGACCGAGATGCTG 370

RESULT 10

ACF72027

ID ACF72027 standard; DNA; 339 BP.

XX AC ACF72027;

DT 20-NOV-2003 (first entry)

DE Photorhabdus luminescens nucleotide sequence #10494.

XX Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;

KW detection; food; gene expression; plant; animal; microorganism; toxin;

KW antibiotic; biopesticide; virulence factor; disease model; plague;

KW whooping cough; gene; ds.

XX Photorhabdus luminescens.

XX WC200294867-A2.

XX 28-NOV-2002.

XX 07-FEB-2002; 2002WO-IB003040.

XX 07-FEB-2001; 2001FR-00001659.

XX (INSP) INST PASTEUR.

XX (CNRS) CNRS CENT NAT RECH SCI.

XX Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;

PI Buchrieser C;

XX WPI; 2003-148459/14.

XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,

PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.

PS Claim 2; SEQ ID NO 10494; 1205pp; French.

XX The invention relates to the isolation of genes and their encoded

CC proteins from Photorhabdus luminescens. The isolated sequences are

CC sources of probes and primers for detecting the genome of P. luminescens

CC and related species; to study polymorphisms; for gene analysis and for

CC detection/amplification of the genes. Antibodies (Ab) raised against the

CC polypeptides encoded by the genes are used for detection/identification

CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that

CC carry a gene-containing vector are used to select compounds that

CC modulate, regulate, induce or inhibit expression of the genes in plants,

CC animals or microorganisms other than P. luminescens and are able to alter

CC response or sensitivity to toxins and antibiotics produced by P.

luminescens. Cells transformed to express the genes are useful for recombinant production of the proteins, particularly toxins and antibacterials useful as insecticides, bactericides and fungicides. The genes, proteins, vectors containing the genes and Ab are also useful therapeutically (to treat microbial infection by bacteria or fungi that are sensitive to P. luminescens-encoded toxins or antibiotics) and as biopesticides. Other uses of the genes and the proteins are as virulence factors and for identifying targets of human diseases for which P. luminescens is a model (particularly plague and whooping cough). This sequence represents one of the isolated P. luminescens genes

SQ Sequence 339 BP; 105 A; 51 C; 96 G; 87 T; 0 U; 0 Other;

Query Match 13.7%; Score 81.4; DB 7; Length 339;
 Best Local Similarity 55.0%; Pred. No. 4.2e-14;
 Matches 182; Conservative 0; Mismatches 146; Indels 3; Gaps 1;

QY 221 ACAAGTGAAGCAATCTCAGGCGCTGCGAGTCTCGCAAGTTCTCGGCTTTCCTAA 280
 Db 5 AAAGATTGATCGGATTATCAACCTTCAATTAGATGATGCGTGAAGCTCTGGCGG 64

QY 281 AAATTGGTATTCGAGGTTACTGTTCTGATGTTGAGGTTTGGTCTCAAGGTGGTT 340
 Db 65 AAGTGGGTATCACCGAATGACAGTAAACAGAGGTGAAGGTTTGGCGCCCAAAAGGTC 124

QY 341 CAACCTGAGAGCGGCGCTCAGAAATTTCTGAGACAAAGTTTCTGTAAGTTTACA 400
 Db 125 ATACAGAGCTGTATCGCGGTGAGAAATATATGTTGAT---TTCTGCCAAAGTGAATA 181

QY 401 TGGAGATCGTGTGTAAGCAAGACCGAGTTGAGGATGTTATAGAAATAATCATTGAGGAGG 460
 Db 182 TAGAATTTGTCGCGCAGATGATATGTCGATACCTGTTGAACCATATGCGACGG 241

QY 461 CAAGACTGAGAGATTGAGACGCGCAGATTTCTTGCTGCTGCTGTTTCAAGTTAATA 520
 Db 242 CACAGACCGGAAATTCGATGATGTAATAATTTGATTTGATGATGACGCTGTTGTC 301

QY 521 GAGTCCGCACTGGTACGCGGCTGATAGGC 551
 Db 302 GTATCCGACCGGTGAGCAGGATGAGAGGC 332

RESULT 11
 ACF67367_52
 Continuation (53 of 57) of ACF67367 from base 5200001 (Photorhabdus luminescens nucleotide sequence split into 57 fragments LOCUS ACF67367 Accession ACF67367

WP	Fragment Name	Begin	End
WP	ACF67367_00	1	110000
WP	ACF67367_01	100001	210000
WP	ACF67367_02	200001	310000
WP	ACF67367_03	300001	410000
WP	ACF67367_04	400001	510000
WP	ACF67367_05	500001	610000
WP	ACF67367_06	600001	710000
WP	ACF67367_07	700001	810000
WP	ACF67367_08	800001	910000
WP	ACF67367_09	900001	1010000
WP	ACF67367_10	1000001	1110000
WP	ACF67367_11	1100001	1210000
WP	ACF67367_12	1200001	1310000
WP	ACF67367_13	1300001	1410000
WP	ACF67367_14	1400001	1510000
WP	ACF67367_15	1500001	1610000
WP	ACF67367_16	1600001	1710000
WP	ACF67367_17	1700001	1810000
WP	ACF67367_18	1800001	1910000
WP	ACF67367_19	1900001	2010000
WP	ACF67367_20	2000001	2110000
WP	ACF67367_21	2100001	2210000
WP	ACF67367_22	2200001	2310000
WP	ACF67367_23	2300001	2410000
WP	ACF67367_24	2400001	2510000
WP	ACF67367_25	2500001	2610000

WP	Fragment Name	Begin	End
WP	ACF67367_26	2600001	2710000
WP	ACF67367_27	2700001	2810000
WP	ACF67367_28	2800001	2910000
WP	ACF67367_29	2900001	3010000
WP	ACF67367_30	3000001	3110000
WP	ACF67367_31	3100001	3210000
WP	ACF67367_32	3200001	3310000
WP	ACF67367_33	3300001	3410000
WP	ACF67367_34	3400001	3510000
WP	ACF67367_35	3500001	3610000
WP	ACF67367_36	3600001	3710000
WP	ACF67367_37	3700001	3810000
WP	ACF67367_38	3800001	3910000
WP	ACF67367_39	3900001	4010000
WP	ACF67367_40	4000001	4110000
WP	ACF67367_41	4100001	4210000
WP	ACF67367_42	4200001	4310000
WP	ACF67367_43	4300001	4410000
WP	ACF67367_44	4400001	4510000
WP	ACF67367_45	4500001	4610000
WP	ACF67367_46	4600001	4710000
WP	ACF67367_47	4700001	4810000
WP	ACF67367_48	4800001	4910000
WP	ACF67367_49	4900001	5010000
WP	ACF67367_50	5000001	5110000
WP	ACF67367_51	5100001	5210000
WP	ACF67367_52	5200001	5310000
WP	ACF67367_53	5300001	5410000
WP	ACF67367_54	5400001	5510000
WP	ACF67367_55	5500001	5610000
WP	ACF67367_56	5600001	5648894

Query Match 13.7%; Score 81.4; DB 7; Length 110000;
 Best Local Similarity 55.0%; Pred. No. 5.5e-13;
 Matches 182; Conservative 0; Mismatches 146; Indels 3; Gaps 1;

QY 221 ACAAGTGAAGCAATCTCAGGCGCTGCGAGTCTCGCAAGTTCTCGGCTTTCCTAA 280
 Db 49307 AAAGATTGATCGGATTATCAACCTTCAATTAGATGATGCGTGAAGCTCTGGCGG 49366

QY 281 AAATTGGTATTCGAGGTTACTGTTCTGATGTTTCCAGGTTTGGTGTCTCAAGGTGGTT 340
 Db 49367 AAGTGGGTATCACCGAATGACAGTAAACAGAGGTGAAGGTTTGGCGCCCAAAAGGTC 49426

QY 341 CAACCTGAGAGCGGCGCTCAGAAATTTCTGAGACAAAGTTTCTGTAAGTTTACA 400
 Db 49427 ATACAGAGCTGTATCGCGGTGAGAAATATATGTTGAT---TTCTGCCAAAGTGAATA 49483

QY 401 TGGAGATCGTGTGTAAGCAAGACCGAGTTGAGGATGTTATAGAAATAATCATTGAGGAGG 460
 Db 49484 TAGAATTTGTCGCGCAGATGATATGTCGATACCTGTTGAACCATATGCGACGG 49543

QY 461 CAAGAACTGAGAGATTGAGACGCGCAGAAATTTCTTGCTGCTGCTGTTTCAAGTTAATA 520
 Db 49544 CACAGACCGGAAATTCGATGATGTAATAATTTGATTTGATGATGACGCTGTTGTC 49603

QY 521 GAGTCCGCACTGGTACGCGGCTGATAGGC 551
 Db 49604 GTATCCGACCGGTGAGCAGGATGAGAGGC 49634

RESULT 12
 ACF65387_3
 Continuation (4 of 7) of ACF65387 from base 300001 (Photorhabdus luminescens nucleotide sequence split into 7 fragments LOCUS ACF65387 Accession ACF65387

WP	Fragment Name	Begin	End
WP	ACF65387_0	1	110000
WP	ACF65387_1	100001	210000
WP	ACF65387_2	200001	310000
WP	ACF65387_3	300001	410000
WP	ACF65387_4	400001	510000
WP	ACF65387_5	500001	610000
WP	ACF65387_6	600001	696798

Query Match 13.7%; Score 81.4; DB 7; Length 110000;
Best Local Similarity 55.0%; Pred. No. 5.5e-13;
Matches 182; Conservative 0; Mismatches 146; Indels 3; Gaps 1;

QY 221 ACAAGTGGAGCAATCTCAGGCGCTCGCAGTCTCGCAAGTTTCTCGGCTTTGCTAA 280
DB 88821 AAAAGATTGATGCGATTATCAAACTTTCAAATTAGATGATGCGTGAAGCTCTGGCGG 88880
QY 281 AAATTGGTATTCGAGGTGTTACTGTTCTGATGTTGAGGTTTGGTCTCAAGTGTT 340
DB 88881 AAGTGGGTATCACCGGATGACAGTAACAGAGGTGAAGGTTTGGGCGCCAAAAGGTC 88940
QY 341 CAATCTGAGGCGAGCGGCTCAGATTTTCTGAAGCAAGTTTGTCTAAAGTTAAGA 400
DB 88941 ATACAGAGCTGATCGCGGTGCAATATATATGTTGGAT---TTTCTGCCAAAGTGAAGA 88997
QY 401 TGGAGATCGTGGTTAGCAAGACCAAGGTTGAGGATGTTATAGAAAATCATTTGAGGAGG 460
DB 88998 TAGAAATTGTCGTCAGATGATATTGTCGATACCTGTGTTGAAACCATTTATGACAGCG 89057
QY 461 CAAGAACTGGAGAGATTGGAGCGGCAAGATTTTCTGCTGCTGTTTTCAGATGTAATA 520
DB 89058 CACAGACCGGAAAATCGTGTATGTTAAATATTTGTTATTTGATGTAGACGTTGTTGCG 89117
QY 521 GAGTCCGCACTGTGAGCGGCTGATAGGC 551
DB 89118 GTATCCGACCGGTGAGCAGGATGAAGAGGC 89148

RESULT 13
AAT42063_03
Confirmation (4 of 19) of AAT42063 from base 300001 (Haemophilus influenzae complete genome)
WP Sequence split into 19 fragments LOCUS AAT42063 Accession Aat42063

WP Fragment Name Begin End
WP AAT42063_00 1 110000
WP AAT42063_01 100001 210000
WP AAT42063_02 200001 310000
WP AAT42063_03 300001 410000
WP AAT42063_04 400001 510000
WP AAT42063_05 500001 610000
WP AAT42063_06 600001 710000
WP AAT42063_07 700001 810000
WP AAT42063_08 800001 910000
WP AAT42063_09 900001 1010000
WP AAT42063_10 1000001 1110000
WP AAT42063_11 1100001 1210000
WP AAT42063_12 1200001 1310000
WP AAT42063_13 1300001 1410000
WP AAT42063_14 1400001 1510000
WP AAT42063_15 1500001 1610000
WP AAT42063_16 1600001 1710000
WP AAT42063_17 1700001 1810000
WP AAT42063_18 1800001 1830121

Query Match 13.1%; Score 77.6; DB 2; Length 110000;
Best Local Similarity 55.1%; Pred. No. 7.7e-12;
Matches 174; Conservative 0; Mismatches 139; Indels 3; Gaps 1;

QY 221 ACAAGTGGAGCAATCTCAGGCGCTCGCAGTCTCGCAAGTTTCTCGGCTTTGCTAA 280
DB 63919 AAAAAATCGAAGCAATGATTAACCCCTTTAAATTAGACGATGTGCGAGAAAGTCTTTGAG 63978
QY 281 AAATTGGTATTCGAGGTGTTACTGTTCTGATGTTGAGGTTTGGTCTCAAGTGTT 340
DB 63979 ATATTGGTATTCAGGTATGACATCAAGAGTACGCGGATTTGGTCTCAAAAGGTC 64038
QY 341 CAATCTGAGGCGAGCGGCTCAGATTTTCTGAAGCAAGTTTGTCTAAAGTTAAGA 400
DB 64039 ATACAGAACTTTATCTGTTGGTGGGATATATATGTTGGAT---TTTCTGCCAAAGTGAAT 64095
QY 401 TGGAGATCGTGGTTAGCAAGACCAAGGTTGAGGATGTTTATAGAAAATCATTTGAGGAGG 460

DB 64096 TCGAAGTGGTGTCTCTCATGAGCTTGTGATCAATGTATTGAAGCGATTATTGAAACGG 64155
QY 461 CAGAACTGGAGAGATTGGAGCGGCAAGATTTTCTGCTGCTGTTTTCAGATGTAATAA 520
DB 64156 CACAAACAGGTAATAATCGTGCAGCGCAAAATTTGTTTATCAGCTTGAGAGGCCATCC 64215
QY 521 GAGTCCGCACTGGTGA 536
DB 64216 GCATTCCGACGGCGA 64231

RESULT 14
ABZ40022
ID ABZ40022 standard; DNA; 363 BP.
AC ABZ40022;
XX 07-MAR-2003 (first entry)
DT N. gonorrhoeae nucleotide sequence SEQ ID 4633.
DE Antibacterial; infection; vaccine; gene therapy; gene; ds.
XX Neisseria gonorrhoeae.
OS Neisseria gonorrhoeae.
PN WC200279243-A2.
XX 10-OCT-2002.
PF 12-FEB-2002; 2002MO-IB002069.
XX 12-FEB-2001; 2001GB-00003424.
PR (CHIR-) CHIRON SPA.
XX Fontana MR, Pizza M, Massignani V, Monaci E;
PI WPI; 2003-058415/05.
DR P-PSDB; ABP79052.
XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a
PT medicament for treating or preventing N. gonorrhoeae infection.
XX Disclosure; Page 528; 815pp; English.
XX The present invention relates to proteins from Neisseria gonorrhoeae.
CC Also disclosed are the nucleic acid molecules encoding the proteins and
CC antibodies that specifically bind to the proteins. The composition
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing N. gonorrhoeae
CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records ABZ37706-ABZ42016 represent nucleic acid
CC molecules of the invention
XX Sequence 363 BP; 84 A; 77 C; 123 G; 79 T; 0 U; 0 Other;

Query Match 13.0%; Score 77; DB 7; Length 363;
Best Local Similarity 54.1%; Pred. No. 9.1e-13;
Matches 180; Conservative 0; Mismatches 150; Indels 3; Gaps 1;

QY 221 ACAAGTGGAGCAATCTCAGGCGCTCGCAGTCTCGCAAGTTTCTCGGCTTTGCTAA 280
DB 32 AAAAAATCGAGCGCATTTGTCAAAACCGTTCAAGCTCGACGCTGCGCGAGGCGTTGACCG 91
QY 281 AAATTGGTATTCGAGGTGTTACTGTTCTGATGTTGAGGTTTGGTCTCAAGTGTT 340
DB 92 AAATCGGCATTTACGGGCGATGACCGTACGAGGATCAAGGGTTCGCGAGGCGAAGGGCG 151
QY 341 CAATCTGAGGCGAGCGGCTCAGAAATTTTCTGAAGCAAGTTTGTCTAAAGTTAAGA 400
DB 152 ATACGGAATCTATCGGCTGCGGATATACCGCTCGAT---TTCTGCCCAAGTCAAA 208
QY 401 TGGAGATCGTGGTTAGCAAGACCAAGGTTGAGGATGTTTATAGAAAAAATCATTTGAGGAGG 460

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 27, 2004, 19:07:55 ; Search time 1376.09 Seconds
(without alignments)
12890.278 Million cell updates/sec

Title: US-09-756-541-16

Perfect score: 594

Sequence: 1 CGGCACGAGGCTACTCGGAA.....CTGCATGAGTACTTCTGCT 594

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:* 1: em_estba.*

2: em_esthum.*

3: em_estcin.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_htc.*

9: gb_estci.*

10: gb_est2.*

11: gb_htc.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_mam.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rod.*

26: em_gss_phg.*

27: em_gss_vrl.*

28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	408.8	68.8	542	14	TI5252 crs852 lamb
2	321	54.0	782	14	CB289081 V-B-112F0
3	300.8	50.6	723	13	BQ798001 EST 6939
4	298.6	50.3	597	10	BE239550 EST403599

5	298.6	50.3	714	12	BI311072
6	297	50.0	812	14	CA922829
7	293	49.3	612	10	BF645522
8	276.6	46.6	759	14	CF920467
9	275.2	46.3	711	12	BI932336
10	275	46.3	732	9	AJ558383
11	274.2	46.2	648	9	AI773079
12	273.6	46.1	613	9	AW035791
13	273.6	46.1	702	12	BI932123
14	273.6	46.1	736	12	BI932913
15	273.6	46.1	741	10	AW738071
16	273.6	46.1	804	12	BI930336
17	273.6	46.1	810	12	BI930060
18	272.6	45.9	799	12	BM113473
19	272	45.8	630	10	AW929534
20	272	45.8	685	12	BI929601
21	270.4	45.5	756	12	BG595705
22	270.2	45.5	677	10	BE823431
23	269.2	45.3	866	13	C81819
24	267	44.9	719	12	BJ573666
25	266.2	44.8	732	14	CD820903
26	265.4	44.7	657	14	CB349285
27	265.4	44.7	701	14	CB349370
28	264.6	44.5	458	10	BE210153
29	264.6	44.5	669	14	CD822497
30	263	44.3	685	14	CD825321
31	263	44.3	745	14	CD819046
32	262.2	44.1	675	14	CB346213
33	248.4	41.8	549	14	CB261070
34	248.2	41.8	560	14	CA514765
35	246.4	41.5	579	10	AW223643
36	241.2	40.6	673	14	CB085731
37	239.2	40.3	540	14	CB349912
38	238.6	40.2	551	14	CB349990
39	238.6	40.2	598	14	CB350294
40	236.4	39.8	540	10	AW624562
41	233.4	39.3	532	14	CB350222
42	227.6	38.3	564	13	BQ118918
43	227.4	38.3	729	14	CF475242
44	227.2	38.2	796	14	CF395143
45	223.6	37.6	690	14	CA290108

ALIGNMENTS

RESULT 1	TI5252	542 bp	mrna	linear	EST 23-MAY-2001
LOCUS	crs852	lambdaZAPST	Ricinus communis	cDNA clone pcrs852	similar to
DEFINITION	nitrogen-regulatory protein, mRNA sequence.				
ACCESSION	TI5252				
VERSION	TI5252.1	GI:14190796			
KEYWORDS	EST.				
SOURCE	Ricinus communis (castor bean)				
ORGANISM	Ricinus communis				
REFERENCE	1 (bases 1 to 542)				
AUTHORS	vandeLoe F.J., Turner, S. and Somerville, C.				
TITLE	Expressed sequence tags from developing castor seeds				
JOURNAL	Plant Physiol. 108, 1141-1150 (1995)				
COMMENT	Contact: Somerville CR Carnegie Institution Carnegie Institution, 290 Panama St, Stanford, CA 94305 Tel: 4153251521 Email: crs@andrew.stanford.edu				
FEATURES	Seq primer: T3.				
source	Location/Qualifiers				
	1..542				
	/organism="Ricinus communis"				

/mol_type="mRNA"
/strain="Baker 296"
/db_xref="taxon:3988"
/clone_lib="pors852"
/clone_lib="lambdazapsr"

/note="Vector: lambdaZAPRII; Site 1: EcoRI; Site 2: XhoI; Poly(A)+ RNA was purified from developing stage III to stage V (Greenwood & Bewley, Can. J. Bot. 60:1751-1760, 1982) endosperm plus embryo of immature castor fruits. cDNA was synthesized and cloned into lambdaZAPRII according to the instructions of the manufacturer (Stratagene); synthesis was primed from the poly(A) tail, and cloned directionally into XhoI (3') and EcoRI (5') sites. In few cases, sequence data indicated that this directionality was reversed. Partial cDNA clones predominate."

ORIGIN

Query Match 68.8%; Score 408.8; DB 14; Length 542;
Best Local Similarity 95.5%; Pred. No. 9.4e-103;
Matches 428; Conservative 0; Mismatches 18; Indels 2; Gaps 1;

```

QY 10 GCTACTGCGAACTGGCTGCTCACTCCTCTTCAATCTAATACATCAAGAAATTC 69
DB 1 GCTACTGCGAACTGGCTGCTCACTCCTCTTCAATCTAATACATCAAGAAATTC 60
QY 70 CCTGTTTTCATTTTCAGTTTGTGTTTCCAGAGCTTAGACATTTCTCGGTTTTCACATT 129
DB 61 CCTGTTTTCATTTTCAGTTTGTGTTTCCAGAGCTTAGACATTTCTCGGTTTTCACATT 120
QY 130 AACACCGCGTCAAGCGGTAAGATATGCCCCCGTTCCTGTGATTAATGCCAAAGC 189
DB 121 AACACCGCGTCAAGCGGTAAGATATGCCCCCGTTCCTGTGATTAATGCCAAAGC 180
QY 190 TCGCCTGACTACATTTCTGTGATGCTTAATCTCAAAAGTGGAGCAATCTCAGGCCCTGG 249
DB 181 TCGCCTGACTACATTTCTGTGATGCTTAATCTCAAAAGTGGAGCAATTTCTCAGGCCCTGG 240
QY 250 CGAGTCTGCGAAAGTTTCTCGGCTTTCCTAAAGTTGTTATTCAGAGTGTACGTTTCT 309
DB 241 CGAGTCTGCGAAAGTTTCTCGGCTTTCCTAAAGTTGTTATTCAGAGTGTACGTTTCT 300
QY 310 GATGTTTCGAGTTTGTGTTCTCAAGTGTGTTCAACTGAGAGCGAGGCGGCTCAGAAATT 369
DB 301 GATGTTTCGAGTTTGTGTTCTCAAGTGTGTTCAACTGAGAGCGAGGCGGCTCAGAAATT 360
QY 370 TCTGAGACAGATTTGTTGCTTAAGTTAAGTGGAGATCGTGTGTTAGCAAGACCGATT 429
DB 361 NCTGAAGNCAGTTTGTGTTCAAGTAAAGATGAGAGCCGCGGT--NGCAAGCCGTT 418
QY 430 GAGGATGTTATAGAAAAATCATTGAGG 457
DB 419 GAGNTGTTATGGAATATTTGGGG 446

```

RESULT 2

CB289081
LOCUS V-B-112F06 VAN-Baker-1 Vitis aestivalis cDNA clone V-B-112F06 5',
mRNA sequence.

CB289081
VERSION CE289081.1 GI:28602822

KEYWORDS EST.

SOURCE Vitis aestivalis

ORGANISM

Buxaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; Vitaceae; Vitis.

1 (bases 1 to 782)

Hou, H.S., Phanikanth, T.V., Kovacs, L., and Qiu, W.P.

Expressed sequence tags of young leaf tissues of a

disease-resistant *Vitis aestivalis* var. Norton

Unpublished (2003)

CONTACT: Wenping Qiu

Department of Fruit Science
Southwest Missouri State University-Mountain Grove
9740 Red Spring Road, Mountain Grove, MO 65711-2999, USA
Tel: 417 926 4105
Fax: 417 926 6646

Email: weq070f@msu.edu
Insert Length: 782 Std Error: 0.00

Plate: VAN-Baker-1-12 row: F column: 06

Seq primer: T3 PRIMER

High quality sequence stop: 782

POLYA-No.

Location/Qualifiers

1. 782

/organism="Vitis aestivalis"

/mol_type="mRNA"

/cultivar="Norton"

/db_xref="taxon:3605"

/clone="V-B-112F06"

/tissue_type="Leaf"

/dev_stage="Young leaf"

/lab_host="X110-Gold E. coli"

/clone_lib="VAN-Baker-1"

/note="Vector: pBluescript II SK (+) Phagemid; Site 1: Xho I; Site 2: EcoR I; VAN-Baker-1 is a cDNA library of Norton grape young leaves (*Vitis. aestivalis* var. Norton). Norton grapevines were grown under normal greenhouse conditions. The cDNA synthesis and library construction was performed according to the instruction manual for pBluescript II XR cDNA library construction kit provided by Stratagene."

ORIGIN

Query Match 54.0%; Score 321; DB 14; Length 782;
Best Local Similarity 75.4%; Pred. No. 2.8e-78;
Matches 399; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

```

QY 62 AAGAAATCCCTGTTTTCAGTTTGTGTTTCCAGAGCTTAGACATTTCTCGGTTT 121
DB 65 AAGAAATCCCTGTTTTCAGTTTGTGTTTCCAGAGCTTAGACATTTCTCGGTTT 124
QY 122 CTCATTTTAAACCGCGGTCAAGCGGTAAAGATATGCCCCCGTTCCTCGTGAATATG 181
DB 125 TCCAGTTTAAATTTGTGCTTAAACCGCTCAAGAGATGATCTGTTCTTCCCATTTGTCAGAG 184
QY 182 CCCAAAGCTGCGAGCTACATTTCTGATGCTAAATTTCTCAAAAGTGAAGCAATTTCTCA 241
DB 185 CTCAGAGCTTCCAGATTTATACCCAGACGCGCAGTTTACAAAGTTGAAGCATCTCTGA 244
QY 242 GGCCTCGCGAGCTTCGCAAGTTTCTCGGCTTTTCTTAAATAATTTGGTATTCGAGTGTGA 301
DB 245 GGCCTCGCGAATCCAGCAGGTTTCTTCCGGCTTTTCTGAAATGCGTATTCGTGGTGTGA 304
QY 302 CTGTTTCTGATGTTTCGAGGTTTTCGTTCTCAAGTGTTCAACTGAGAGCGAGGCGGCT 361
DB 305 CTGTTTCTGATGTTTCGAGGCTTTTGGTCTCAAGCGGATCGCTGAAAGACAGGCGGCT 364
QY 362 CAGAAATTTCTGAAGACAAAGTTTCTGCTAAAGTTTAAAGATCGAGATCGTGTGTAGCAAG 421
DB 365 CTGAATTTTCTGAGGACAAATTTGTTGCTTAAAGTTAAATGAGATTTGTTGTGAGCAAG 424
QY 422 ACCAGGTTGAGGATGTTTATAGAAAAATTCATGAGGAGCGCAAGAACTGGAGAGATTGGAG 481
DB 425 ACCAGGTTGAGGAGTAAATTGACAAGATCAATGAGGTGGCAAGGACTGGAGAGATTGGTG 484
QY 482 ACCGCAAGATTTTCTGCTGCTGTTTTCAGATGTAATAAGATTCGCACTGTTGAGGCGGG 541
DB 485 ATGGCAAGATCTTTTTCGTCGCCATATCGGATGTGATAAGAGTTTCGGAACCGGTGAACGTG 544
QY 542 GTGATAAGGCTGAGAGGATGACAGAGGCGGATCTGACATCAGTACTTTC 590
DB 545 GAGAGATGGCTGAGAGGATGACTGTGTGGGCGAATGATATGTTGTCCTACTAC 593

```

RESULT 3

BQ798001/c
 LOCUS BQ798001 723 bp mRNA linear EST 30-JUL-2002
 DEFINITION EST 6939 Ripening Grape berries Lambda Zap II Library Vitis
 vinifera cDNA clone RT093C02 3', mRNA sequence.
 ACCESSION BQ798001
 VERSION BQ798001.1 GI:22012967
 KEYWORDS EST.
 SOURCE Vitis vinifera
 ORGANISM Vitis vinifera
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; Vitaceae; Vitis.
 1 (bases 1 to 723)
 Abbal, P., Agasse, A., Ageorges, A., Atanaseva, R., Barrieu, F.,
 Couture, C., Dedaldecamp, F., Delrot, S., Glissant, D., Grimplet, J.,
 Hamdi, S., Romieu, C. and Terrier, N.
 Generation of Expressed Sequence Tag from Grape Berry (skin, pulp
 or seeds) at Various Developmental Stages
 Unpublished (2002)
 Contact: Romieu C.
 Unite de Recherche des Produits de la Vigne
 Institut National de la Recherche Agronomique
 2, place Viala, 34 060 Montpellier Cedex 01, France
 Tel: 00-33-(0)4-99-61-28-62
 Fax: 00-33-(0)4-99-61-28-57
 Email: romieu@sensam.inra.fr
 Seq primer: 17.
 Location/Qualifiers
 1..723
 /organism="Vitis vinifera"
 /mol_type="mRNA"
 /cultivar="Shiraz"
 /db_xref="taxon:29760"
 /clone="RT093C02"
 /dev_stage="ripening stage"
 /clone_lib="Ripening Grape berries Lambda Zap II Library"
 /note="Organ: Fruit; Vector: Lambda Zap II; Site 1: Eco
 RI; Site 2: XhoI; Oriented library, construction described
 in Generation of ESTs from grape Berry (skin, pulp or
 seeds) at various developmental stages by Terrier, N.,
 Ageorges, A., Abbal, P., Romieu, C. in J. Plant Physiol. 158
 (12): 1575-83 2001"
 ORIGIN
 Query Match 50.6%; Score 300.8; DB 13; Length 723;
 Best Local Similarity 76.0%; Pred. No. 1.1e-72;
 Matches 371; Conservative 0; Mismatches 117; Indels 0; Gaps 0;
 103 CTTAGACATCTCGTTTTCTCATTAAACCGCGGTCAAGCGGTAAAGATATGCCCCC 162
 723 CTCAAAGATCTCGAAATTTCCAGTTTAAATTTGTGCTTAAACCGCTCAAAGAATCGATCT 664
 163 GTGCTTCGTGATTAATGCCCAAGCTCGCTGACTACATCTCTGATCTAAATTTCTAC 222
 663 GTTCTTCCATGTCAGAGCTCAGAGCTCTCCAGATTATACCCAGCGCGAGTTTAC 604
 223 AAAGTGAAGCAATTTCTCAGCCCTCGCGAGTCTCGCAAGTTTCTCGCTTTGCTTAA 282
 603 AAAGTGAAGCGATCTCAGCCCTCGCGAATCGAGAGTTTCTCGCTTTGCTTAA 544
 283 ATTGGTATCGAGGTTTACTGTCTCATGTTTCAGGTTTGTGTCCTCAAGGTTTCA 342
 543 ATGGGTATTCGTGTGTACTGTCTCATGTTTCAGGTTTGTGTCCTCAAGGTTTTCG 484
 343 ACTGAGAGCGCGCTCAGATTTCTCAGACAGATTTGTGTCCTCAAGTTAAGATG 402
 483 CCTGAAGACAGGCGGCTCTGAAATTTTCTGAGGACAAATTTGTGCTTAAAGTCA 424
 403 GAGATCGTGGTTAGCAAGACAGGTTTGAGGATGTTATAGAAAAATCAATTGAGGAGCA 462
 423 GAGATTGTTGGTACAGAACACAGGTTGAACAGATTAATTGACAGATCAATGAGTGGCA 364
 463 AGAATCGAGAGATTGGAGACGGCAAGATTTTCTTGCTGCCTGTTTCAGATGTAATAAGA 522

363 AGGACTGGAGAGATTGGTGTATGCGCAAGATCTTTTGGTGCCCATATCGGATGTGTAAGA 304
 523 GTCCGACTGGTGTAGCGGGGTGATATAGCTGAGAGATGACAGAGCGGCGATCTGACATG 582
 303 GTTCGACCGGTGACGTGGAGAGCTGGCTGAGAGATGACGGGTGGCGCAACTGATATG 244
 583 AGTACTTC 590
 243 TCCACTAC 236
 BE239550 597 bp mRNA linear EST 12-JUL-2000
 EST403559 MHRP- Medicago truncatula cDNA clone pMHRP-28F3, mRNA
 sequence.
 BE239550
 BE239550.1 GI:9055733
 EST.
 Medicago truncatula (barrel medic)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.
 1 (bases 1 to 597)
 Harrison, M.J., Town, C.D., Bowman, C.L., Craven, M.B., Harsen, T.S.,
 Holt, I.E., Cho, J. and Fraser, C.M.
 ESTs from phosphate-starved roots of Medicago truncatula
 Unpublished (2000)
 Contact: Maria J. Harrison
 Plant Biology Division
 The Samuel Roberts Noble Foundation
 2510 Sam Noble Parkway, Ardmore, OK 73401, USA
 Tel: 580-223-5810
 Fax: 580-221-7380
 Email: mjharrison@noble.org
 The Samuel Noble Roberts Foundation: N265417e
 TIGR sequence name: MTHAC26TK
 More information is available at: .
 http://chrysis.tamu.edu/medicago
 Seq primer: SKmod (CTA gAA CTA gGg GAT CC).
 Location/Qualifiers
 1..597
 /organism="Medicago truncatula"
 /mol_type="mRNA"
 /cultivar="A17"
 /db_xref="taxon:3880"
 /clone="pMHRP-28F3"
 /tissue_type="roots"
 /dev_stage="phosphate-starved"
 /lab_host="XL0UR"
 /clone_lib="MHRP-"
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
 XhoI; At the trifoliolate stage, M. truncatula plants were
 transplanted to phosphate-free sand and grown for a
 further 30 days. During this period, they were fertilized
 twice weekly with 1/2 Hoaglands solutions containing 20uM
 potassium phosphate. cDNA was prepared from polyA+
 enriched RNA. The cDNA was directionally ligated into the
 Unizap XR vector from Stratagene and packaged using
 Gigapack III Gold packaging extracts. Plasmids containing
 cDNA inserts were excised from the recombinant lambda-Zap
 phage using Ex-assist helper phage and propagated in
 XL0UR cells."
 ORIGIN
 Query Match 50.3%; Score 298.6; DB 10; Length 597;
 Best Local Similarity 71.8%; Pred. No. 4.6e-72;
 Matches 407; Conservative 0; Mismatches 154; Indels 6; Gaps 1;
 13 ACTGCGAACTGGGCTTGCTCACTCTTCTCATCTTAATAACATCAAGAAAGATTCCCT 72

[illegible]

RESULT 5
 BI311072
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 MEDICAGO TRUNCATULA (barrel medic)
 MEDICAGO TRUNCATULA
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.
 REFERENCE
 1 (bases 1 to 714)
 Grusak, W.A., Samac, D.A., Town, C.D., Van Aken, S., Utterback, T.,
 Cho, J. and Fraser, C.M.
 ESTs from developing reproductive tissues of Medicago truncatula
 Unpublished (2001)
 CONTACT: Michael A. Grusak
 USDA/ARS Children's Nutrition Research Center
 Baylor College of Medicine
 1100 Bates Street, Houston, TX 77030-2600, USA
 Tel: 713-798-7044
 Fax: 713-798-7078
 Email: mgrusak@bcm.tmc.edu
 B397927e
 TIGR sequence name: MTPAP50TK
 More information is available at: www.medicago.org
 Seq primer: SKnod (CTA GAA CTA gtg GAT CCI).
 Location/Qualifiers
 1..714
 /organism="Medicago truncatula"
 FEATURES
 source

ACCESSION CA922829
 VERSION CA922829.1 GI:27409759
 KEYWORDS EST.
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
 REFERENCE 1 (bases 1 to 812)
 AUTHORS VandenBosch, K., Endre, G., Silverstein, K., Town, C.D., Van Aken, S., Utterback, T., Cheung, F., and Fraser, C.M.
 TITLE The Medicago truncatula 6K unigene set: cDNA clones selected and re-arrayed from various libraries
 JOURNAL Unpublished (2002)
 COMMENT Contact: VandenBosch K
 Department of Plant Biology
 University of Minnesota
 220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
 Tel: 612 624 2755
 Fax: 612 625 1738
 Email: kvandenb@chs.umn.edu
 Alias Clone pMHP-28f3
 TIGR sequence name: MTUCF91TV
 More information is available at: www.medicago.org
 Seq primer: (gtA AtA CGA Ctc Act AtA ggg C).
 FEATURES
 source
 1..812
 /organism="Medicago truncatula"
 /mol_type="mRNA"
 /cultivar="Al7"
 /db_xref="taxon:3880"
 /clone="MTUS-58H7"
 /tissue="mixed tissues"
 /dev_stage="various stages"
 /lab_host="XL0LR"
 /clone_lib="MTUS"
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XL0LR cells."
 ORIGIN
 Query Match 50.0%; Score 297; DB 14; Length 812;
 Best Local Similarity 71.6%; Pred. No. 1.3e-71;
 Matches 406; Conservative 0; Mismatches 155; Indels 6; Gaps 1;
 13 ACTCGAACTGGGCTTGTCTCACTCTCTTCAATCTTAATCAATCAAGAAAGATTCCCT 72
 768 ATTGGAACCGAACGGTTCACGGTTGAATTTTCATATCAATGAATCAATTTCT 709
 73 GTTTTGAATTCAGTTGTTTGTCCAGAGCTTAGACATTCCTCGGTTTCTCACTTAAAC 132
 708 TTTTCAAGCTTCAGGTCATTCGAAGGTTTCGGAGATTC-----TTCTCATCGCAAT 655
 133 ACCGGGTCAAGCGGTAAAGATATGCCCGCTCGTCTCTGTGATTAATGCCAAGCTCG 192
 654 GTGGTCTTAAAGAACACGGAAATGCATCGATTCTTCCCAAAATCAGAGCTCAAAACCTT 595
 193 CCTGACTACATTCCTGATGTCTAAATTTCAAAAGTGAAGCAATTCAGAGCCCTGGCGA 252
 594 CCTGACTATGTTCTGATCCCAAGTTTACAAAGTTGAAGCAATTCAGGCCATGGAGA 535
 253 GTCTCGAAGTTTCTCGGCTTTGCTAAAAATGGTATTCAGAGGTGTTACTGTTCTCAT 312
 534 ATCCCTCAGGTTTCTTCGGGTTTGTGAATGGAAATGCGTGTCTCATGTATCTGAT 475
 313 GTTCAGGTTTGTGCTCAGAGGTGGTTCACTAGAGGCGAGGCGCTCAGAAATTTCT 372
 474 GTCAAGGGGTTTGTGCTCAGGGTGGCTCAAAAGAGAGCGGAGGCTCCGAATTTCT 415

QY 373 GAAGACAAGTTTCTGCTAAGTTAAGATCGAGATCGTGTACCAAGACACAGTTGAG 432
 Db |||||
 414 GAAGACAATTTTGTGTCAAAAGTTAAATAGTGTGAGAAAAGACCAAGTTGAG 355
 QY 433 GATGTTATAGAAAATATCATTTGAGAGCGCAAGACTCGAGAGATTGCGAGCGCAAGATT 492
 Db |||||
 354 GCAGTGATAACCAATTTATGGAGACGGCAAGAACTGGGGAGATTGTTGATGGCAAAATT 295
 QY 493 TTCTTGCTGCTGTTTTCAGATGTAATAAGAGTCCGCACTGCTGGTGGCGGGTGATAAGGCT 552
 Db |||||
 294 TTCTTGATCCCTGTATCTGATGTAATAAGAAATCCGACAGGTGAGCGTGGGGAGCAGGCT 235
 QY 553 GAGAGGATCAGAGGAGGGGATCTGAC 579
 Db |||||
 234 GAGGAGATGCTGGGGGACTAACTGAC 208

RESULT 7
 BF645522 612 bp mRNA linear EST 20-DEC-2000
 LOCUS clone NF036F08EC 5', mRNA sequence.
 DEFINITION BF645522
 ACCESSION BF645522
 VERSION BF645522.1 GI:11910651
 KEYWORDS EST.
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula

REFERENCE 1 (bases 1 to 612)
 AUTHORS Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.
 TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation - Center for Medicago Genomics Research
 JOURNAL Unpublished (2000)
 COMMENT Contact: Bixon RA
 Plant Biology Division
 The Samuel Roberts Noble Foundation
 2510 Sam Noble Parkway, Ardmore, OK 73402, USA
 Tel: 580 221 7302
 Fax: 580 221 7380
 Email: radixon@noble.org
 Insert Length: 612 Std Error: 0.00
 Plate: 036 row: F column: 08
 Seq primer: TCACACAGGAAACAGCTATGAC.

FEATURES
 source
 1..612
 /organism="Medicago truncatula"
 /mol_type="mRNA"
 /db_xref="taxon:3880"
 /clone="NF036F08EC"
 /tissue_type="Cell suspensions derived from root tissues"
 /dev_stage="Cell suspensions were subcultured every 14 days. Cells were induced six days after subculture"
 /clone_lib="Elicited cell culture"
 /note="Vector: Lambda Zap; Cells were induced with yeast cell wall extracts equivalent to 50ug/ml glucose in the final concentration. Samples were taken at 0.5, 1, 12 and 24 hours after induction. Equal amounts of RNA from each time point were pooled and used for mRNA isolation."

ORIGIN
 Query Match 49.3%; Score 293; DB 10; Length 612;
 Best Local Similarity 70.9%; Pred. No. 1.7e-70;
 Matches 402; Conservative 0; Mismatches 159; Indels 6; Gaps 1;
 13 ACTGGAACCTGGGCTTGTCTCACTCTCTTCAATCTTAATCAATCAAGAAAGATTCCCT 72
 44 ATTGGAACCGAAGCTGTTCAACGGTTTGAATTTTCATATCAATCAATCAATTTCT 103

73 GTTTTTCATTTTCAGTTGTTTTCAGAGCTTAGACATTTCTCGTTTCTCACTTTAAC 132
 104 TTTTCAAGCTTCAGCTTCATTCGAAGCTTTCGAGNTTC-----TTCTCATCGCAAT 157
 133 ACCGCGTCAAGCGCGTAAGATATGCCCCGCTGCTTCTGTGATTAATGCCCAAAGCTCG 192
 158 GTGTCCTAAAGCAACGGAATGCAATCTTCTCCAAAATCAGAGCTCAAAAACCTT 217
 193 CCTGACTACATTCCTGATGCTAAATTTTACAAAGTGAAGCAATTTCTCAGGCGCTGGGA 252
 218 CCGTACTATGTTCTGAATCCAGTTTACAAAGTTGAAGCCATTTCTCAGGCCATGAGA 277
 253 GTCTCGCAAGTTTCTCGCTTTCGTAATAATTCGATTTTCAGGTTTCTGATTTCTGAT 312
 278 ATCCCTCAGTTTCTTCGGTTTGTGAAATGGAATTCGTGTGCTACATTTCTGAT 337
 313 GTTCGAGTTTGTGCTCAAGTGGTTCAACTGAGAGCGAGCGGCTCAGAAATTTTCT 372
 338 GTCAAGGGGTTTGTGCTCAGGTTGGCTCAAAAGATAGGCGAGGAGGCTCCGAATTTTCT 397
 373 GAAGCAAGTTTGTGCTTAAAGTTAAAGTGGAGATCGTTGTTAGCAAGACAGGTTGAG 432
 398 GAAGCAATTTTGTGCTAAAGTTAAAGTGGAGATAGTGGTGAAGAAAGACAGGTTGAG 457
 433 CATGTTATAGAAAATCATTGAGAGCGAGCAACTGGAGAGATTGGAGAGCGCAAGATT 492
 458 GCAGTGATAACAAAATTTTCGACACGACAGACTGGGAGATTGGTATGGCAAAATT 517
 493 TTCTGTGCTGCTTTCAGATGTAATAGAGTCCGCACTGGTGAAGCGGGGTGATAAGCT 552
 518 TTCTGTATCCTGTATCTGATGTAATAAGAAATCCNACAGTGCAGCGTGGGGGAGCANGCT 577
 553 GAGAGGATGACAGAGCGGCGATCTGAC 579
 578 GANAGGATGCTGGGGGACTAACTGAC 604

RESULT 8
 CF920467
 LOCUS gmhrRw3-14_G03_1_019 Soybean root hair subtracted cDNA library
 DEFINITION gmhrRw3 Glycine max cDNA, mRNA sequence.
 ACCESSION CF920467
 VERSION
 KEYWORDS EST.
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max

REFERENCE
 AUTHORS Scheffler,B.E., Huang,S., Liu,X., Nguyen,H., Duke,M. and Stacey,G.
 TITLE Expressed sequence tags from soybean root hair subtractive cDNA library
 JOURNAL Unpublished (2003)
 COMMENT Contact: Gary Stacey
 University of Missouri
 108 Waters Hall, Columbia, MO 65211, USA
 Tel: 573-884-4752
 Fax: 573-882-0588
 Email: staceyg@missouri.edu
 Single pass sequence
 Seq primer: T7.

FEATURES
 source
 1. .759
 /location/Qualifiers
 /organism="Glycine max"
 /mol_type="mRNA"
 /cultivar="Williams 82"
 /db_xref="taxon:3847"
 /tissue_type="root hairs"
 /clone_lib="Soybean root hair subtracted cDNA library
 gmhrRw3"

ORIGIN

Query Match 46.6%; Score 276.6; DB 14; Length 759;
 Best Local Similarity 73.9%; Pred. No. 6.3e-66;
 Matches 351; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 118 TTTTCTGCTTTTAAACCGCGTCAAGCGCTGAAGATATGCCCCGCTGTTCTCTGATTT 177
 Db 127 TCTCTCAGCGCAATGTGGCTCTAAGACGACAGATAAATGGAACAATTTCTTCCCAATC 186
 QY 178 AATGCCAAAGCTCGCTGACTAGATCTCTGATGCTAAATTTCTACAAAGTGAAGCAAT 237
 Db 187 AGAGCCCAAGATCTTCCAGATTTATGTTCCAAATCCGAGTTTACAAAGTAGAGCGATT 246
 QY 238 CTCAGGCCCTCGGAGTCTCGCAAGTTTCTCGGCTTTGCTTAAATAATGGTATTCGAGT 297
 Db 247 CTCAGGCCATGGCGAGTTTCCCGAGTTTCTCGGGCTTTGTTGAAATGGAATTCGTGCT 306
 QY 298 GTTACTGTTTCTGATGTTTCGAGGTTTGTGCTCAAGTGGTTCAACTCAGAGGCGAGGC 357
 Db 307 GTCACTGTATCTGATGTCAGGGGCTTTGGTGTCTCAGGTTGGTTCAAAAGAGAGGCGAG 366
 QY 358 GGCTCAGAAATTTCTGAAGACAAAGTTTGTGCTAAAGTTAAGATGAGATCGTGTATGAC 417
 Db 367 GGCTCCGAAATTTTCAGAGACAAATTTGTTCCAAAGTTAAATGGAAGTAGTGTGAGA 426
 QY 418 AAAGACCAAGTTTGAAGATGTTTATAGAAAATTCATTTGAGAGGCGAAGCACTGGAGATT 477
 Db 427 AAGGACCAAGTTGAGGCGAGTTTATGACAAAATTTATGAGAGGCGAAGCACTGGGAGATT 486
 QY 478 GGAGAGCGCAGATTTTCTGCTGCTGTTTCAGATGTAATAAGATCGGCACTGGTGA 537
 Db 487 GGTGATGSCAAAATTTTCTGATCCCATCTCAGATGTTATAAGAAATTCGACAGGTGAA 546
 QY 538 CGGGGTGATTAAGGCTGAGAGGATGACAGAGGCGGCTCTGACATGAGTACTTCTG 592
 Db 547 GGTGGGAGCAGGCGAGCTAGATGACTGGGGCGGAGTGCATGTTATCTGCTG 601

BI932336 711 bp mRNA linear EST 18-OCT-2001
 EST552225 tomato flower, 8 mm to preanthesis buds Lycopersicon
 esculentum cDNA clone cF022H10 5' end, mRNA sequence.

RESULT 9
 BI932336
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Lycopersicon esculentum (tomato)
 Lycopersicon esculentum
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 711)

REFERENCE
 AUTHORS van der Hoeven,R.S., Bezzerides,J.L., Karamnycheva,S.A., Tsai,J.,
 Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,
 Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
 TITLE Generation of ESTs from tomato flower tissue, buds 8 mm -
 preanthesis

JOURNAL
 COMMENT Unpublished (2001)
 Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 This clone is available through the Clemson University Genomics
 Institute

Seq primer: T3.
 Location/Qualifiers
 1. .711
 /organism="Lycopersicon esculentum"

/note="Organ: root hairs; Vector: pCR2-1 Topo; cDNA clones
 generated from soybean root hair tissue treated with
 Bradyrhizobium japonicum for 3 hours."

/mol_type="mRNA"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="cTOC22H10"
 /issue_type="flower"
 /dev_stage="buds 8mm to preanthesis buds"
 /clone_lib="tomato flower, 8 mm to Site 1: EcoRI; Site 2:
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
 XhoI; supplier: Cornell University; sequencing: The
 Institute for Genomic Research; Flower buds and flowers
 were taken from greenhouse plants (4-8 wks old, TA496).
 They were immediately frozen in liquid nitrogen and then
 size-separated while remaining frozen."

ORIGIN

Query Match 46.3%; Score 275.2; DB 12; Length 711;
 Best Local Similarity 75.9%; Pred. No. 1.5e-65;
 Matches 340; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 135 CGCGGTCAAGCGGTGAAGATATCCCGCGTCTCTGTGATTAAATGCCCAAGCTCGCC 194
 DB 175 CACCTTCAACGCTGTCAAAATGCTCTTCTCCGATTATCAGAGCCCAAACTCTCC 234
 QY 195 TGACTACATTCCTGATGCTAAATCTACAAAGTGGAGCAATCTCAGGCCCTGGGAGT 254
 DB 235 AGATTTCGTCCTGATGCGAGTGTTCACAAAGTTGAGCAATCTAAGACCTTGGAGAT 294
 QY 255 CTCGGAAGTTCTCTCGCTTTGTCTAAATGTGTAATTCGAGGTGTACTCTTTCTGATGT 314
 DB 295 TCAACAGGTTCTCTCGGCACTACTGAAATGGGCAATTCGTTGCTCACTGTTTCAGATGT 354
 QY 315 TCGAGGTTTGTGCTCAAGTGTGTTCACTGAGAGCGAGGGCTTCAGAAATTTCTGA 374
 DB 355 TCGTGTGTTTGGGCCCAAGTGGCTTGAAGAGGCAAGCTGGCTCTGAATTCCTGA 414
 QY 375 AGACAAAGTTTGTGCTAAAGTTAAGATGGAGATCGTGTGTAGCAAGACCAAGTTTGAGGA 434
 DB 415 AGACAGTTTGTGCAAAAGTTAAATGGAATTTGTGTCAGCAAGACCAAGTTTGAGG 474
 QY 435 TGTATAGAAAATCATTGAGAGGCGAAGAACTGGAGATTTGGAGACCGCAAGATTTT 494
 DB 475 AGTCATTGCCATGATAATTGAAGAGGCAAGAACTGGTGAATAGGTGATGGAAGATAT 534
 QY 495 CTGCTGCTGTTTTCAGATGTATAGAGTCCGACCTGTTGAGCGGGGTGATAGGCTGA 554
 DB 535 CTGACTCCCATCTCCGATGTTATAGAGTTCCGACTGGTGAACGGGAGATAGGCGAGA 594
 QY 555 GAGGATGACAGGAGGCGGATCTGACATG 582
 DB 595 GAGGATGATGGAGGCGCATCTGATG 622

RESULT 10
 AJ558383
 LOCUS
 DEFINITION
 AJ558383 Antirrhinum majus whole plant Antirrhinum majus cDNA clone
 018_1_07_f22, mRNA sequence.
 ACCESSION
 AJ558383
 VERSION
 AJ558383.1 GI:31660955
 KEYWORDS
 EST.
 SOURCE
 Antirrhinum majus (snapdragon)
 ORGANISM
 Antirrhinum majus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; Lamiales; Antirrhinaceae; Antirrhineae;
 Antirrhinum.
 REFERENCE
 1 (bases 1 to 732)
 Zachgo, S., Stueber, K., Saedler, H., Sommer, H. and Schwarz-Sommer, Z.
 Antirrhinum EST collection
 Unpublished (2003)
 JOURNAL
 Contact: Schwarz-Sommer Z
 Molekulare Pflanzen-genetik
 MEZ fuer Zuechtungsforschung

FEATURES
 source
 Location/Qualifiers
 1..732

/organism="Antirrhinum majus"
 /mol_type="mRNA"
 /db_xref="taxon:4151"
 /clone="018_1_07_f22"
 /issue_type="whole plant"
 /clone_lib="Antirrhinum majus whole plant"

ORIGIN

Query Match 46.3%; Score 275; DB 9; Length 732;
 Best Local Similarity 77.8%; Pred. No. 1.7e-65;
 Matches 332; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

QY 166 GTTCTCTGTGATTAAATGCCCAAGCTCGCTGACTATATCTCTGATGCTAAATTTCTACAA 225
 DB 169 GTTCCAATTTGTTAGGGCTCAGATTTCTCCAGAATCCCTACTGACTCCAAGTTCTACAA 228
 QY 226 GTGAGCAATTTCTCAGGCCCTCGGAGTCTCGCAAGTTTCTCGGCTTTGCTAAATAAT 285
 DB 229 GTCGAAGCTATTTAAGACCTTGAGGATCCACAGGTTTCTCGGCTTCTGCTGAAATG 288
 QY 286 GGTATTCGAGGTGTACTGTCTGATGTTTCGAGGTTTGGTGTCTCAAGTGTGTTCAACT 345
 DB 289 GGAATTCGTGTTCACTGTCTGATGATGACGAGGTTTGGTCTCAGGGTGGTTTGACA 348
 QY 346 GAGAGGAGGGGGCTCAGAAATTTCTGAGACAAGTTTGTGTTGCTAAAGTTAAGATGGAG 405
 DB 349 GAAAGACAGTTTGGCTCTGAATTTCTCAGAAACAAATTTGTTGCAAAAGTTAAGATGGAG 408
 QY 406 ATCGTGTGTAGCAAGACCAAGTTTGAGGATGTTATAGAAAATCATTGAGAGGCAAGA 465
 DB 409 ATTGTTGTGAGCAAGACCAAGTTTGATGCGTATAGAAAGATATAGAGGAGCAAGA 468
 QY 466 ACTGGACAGATTGGAGACGGCAAGATTTCTGTGCTGCTGTTTCAGATGTAATAAGAGTC 525
 DB 469 ACAGGAGATTGGTGTGAGGAAATTTTGTGCTTCCAGTTTCTGAATGTAATAAGAGTT 528
 QY 526 CGCACTGTTGAGCGGGGTGATAGGCTGAGAGATGACAGAGGGCGCATCTGATGATGAT 585
 DB 529 CGCACTGTTGAACGAGGGGAAAGGCTGAGAGATGTTGGGAGGACGGGCGCATGTTTC 588
 QY 586 ACTCTCG 592
 DB 589 TCTAATG 595

RESULT 11

AJ773079
 LOCUS
 DEFINITION
 EST254179 tomato resistant, Cornell Lycopersicon esculentum cDNA
 clone CLER5E12, mRNA sequence.
 ACCESSION
 AJ773079
 VERSION
 AJ773079.1 GI:5271120
 KEYWORDS
 EST.
 SOURCE
 Lycopersicon esculentum (tomato)
 ORGANISM
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
 REFERENCE
 1 (bases 1 to 648)
 D'Ascenzo, M., He, X., Lyman, J., Matern, A.L., Vision, T., Holt, I.E.,
 Liang, F., Upton, J., Ronning, C.M., Craven, M.B., Fujii, C.Y.,
 Bowman, C.L., Nierman, W., Fraser, C.M., Venter, J.C., Tanksley, S.D.,
 Giovannoni, J.J. and Martin, G.B.
 Generation of ESTs from Pseudomonas resistant tomato
 Unpublished (1999)
 CONTACT: CUGI
 JOURNAL
 COMMENT
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>

```

5 prime sequence.
FEATURES             source
  1..648
    /organism="Lycopersicon esculentum"
    /mol_type="mRNA"
    /cultivar="R11-12 (35S::Pto in Rio Grande x Money Maker)"
    /db_xref="taxon:4081"
    /clone="CLER5E12"
    /tissue_type="leaf"
    /dev_stage="4-week old"
    /lab_host="SOLR"
    /clone_lib="tomato resistant, Cornell"
    /note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; CLER - Tomato Pseudomonas Resistant EST Library.
Directionally cloned cDNAs inserted into pBlueScript
SK(-) at 5' end with EcoRI and 3' end with XhoI site."

ORIGIN
Query Match      46.2%; Score 274.2; DB 9; Length 648;
Best Local Similarity 75.7%; Pred. No. 2.9e-65;
Matches 339; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 135 CCGGTCAGCGCGTAAGATATGCCCCCGTCTCTGTGATTAATGCCAAGCTCGCC 194
Db 181 CACCTTCAACGCTGTCAAAATGCTCTTTTCCCGATTATCAGAGCCCAAACTCTCC 240

QY 195 TGACTACATTCCTGATGCTAAATTCACAAAGTGAAGCAATTCACAGCCCTGGCGAGT 254
Db 241 AGATTTGCTCCCTGATGGGAAGTTTACAAAGTTGAAGCAATTCGAAGACCTTGAGAAAT 300

QY 255 CTCGCAAGTTTCTCGGCTTTCCTAAAAATTTGATTCGAGGTGTTACTGTTTCTGATGT 314
Db 301 TCACAGGTTTCTTCGGCACACTGAATGGCAATCGTGTGTGTCACCTGTTTCGAGTGT 360

QY 315 TCGAGGTTTGGTCTCAAGGTGTTCAACTGAGAGGCGAGCGGCTCAGAAATTTTCTGA 374
Db 361 TCGTGGTTTGGCGCCCAAGGTGGCTTCACTGAGAGGCAAGCTGGCTCTGAAATTTCTGA 420

QY 375 AGACAAGTTTGTCTAAAGTTAAAGTGAAGTCTGTTAGCAAGACCAAGCTTGAGCA 434
Db 421 AGACACGTTTGTGCAAAAGTTAAATGGAATTTGTTGTCAGCAAGACCAAGCTTGAAGG 480

QY 435 TGTTATAGAAAATTCATTGAGAGGCAAGAACTGGAGAGATTGGAGAGCGGCAAGATTTT 494
Db 481 AGTCATTGCCAAGATAATTGAAGAGCAAGACTGTTGTAATAGTGTGGAAGATATT 540

QY 495 CTTCGTCCTGTTTCAGATGTAATTAAGATCGCACTCGTGTGAGCGGGGTGATAAGGCTGA 554
Db 541 CTTCGTCCTGTTTCAGATGTAATTAAGATCGCACTCGTGTGAGCGGGGTGATAAGGCTGA 600

QY 555 GAGGATGACAGAGGCGGATCTGACATG 582
Db 601 GAGGATGATGGAGGGCATGCTGATATG 628

RESULT 12
AW035791
LOCUS EST281945 tomato callus, TAMU Lycopersicon esculentum cDNA clone
DEFINITION CLRC36D8, mRNA sequence.
ACCESSION AW035791
VERSION AW035791.1 GI:5894547
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 613)
Alcala, J., Vrebalov, J., White, R., Matern, A.L., Vision, T.,
Holt, I.E., Liang, F., Upton, J., Craven, M.B., Bowman, C.D., Ahn, S.,
Ronning, C.M., Fraser, C.N., Martin, G.B., Tanksley, S.D. and
Giovannoni, J.

TITLE Generation of ESTs from tomato callus tissue
JOURNAL Unpublished (1999)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
FEATURES             Location/Qualifiers
  source
    1..613
      /organism="Lycopersicon esculentum"
      /mol_type="mRNA"
      /cultivar="TA496"
      /db_xref="taxon:4081"
      /clone="CLRC36D8"
      /tissue_type="callus"
      /dev_stage="25-40 days old"
      /lab_host="XLI-Blue MRF"
      /clone_lib="tomato callus, TAMU"
      /note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Giovannoni laboratory; CLRC - Coryledons
of seedlings 7-10 days post-germination were excised, cut
at both ends and placed on MS medium with no selection.
Mixed callus was harvested at 25 and 40 days and included
undifferentiated masses. Tomato Callus EST Library"

ORIGIN
Query Match      46.1%; Score 273.6; DB 9; Length 613;
Best Local Similarity 75.7%; Pred. No. 4.2e-65;
Matches 339; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 135 CCGGTCAGCGCGTAAGATATGCCCCCGTCTCTGTGATTAATGCCAAGCTCGCC 194
Db 134 CACCTTCAACGCTGTCAAAATGCTCTCTTTCCGATTATCAGAGCCCAAACTCTCC 193

QY 195 TGACTACATTCCTGATGCTAAATTCACAAAGTGAAGCAATTCACAGCCCTGGCGAGT 254
Db 194 AGATTTGCTCCCTGATGGGAAGTTTACAAAGTTGAAGCAATTCGAAGACCTTGAGAAAT 253

QY 255 CTCGCAAGTTTCTCGGCTTTCCTAAAAATTTGATTCGAGGTGTTACTGTTTCTGATGT 314
Db 254 TCACACGTTTCTTCGGCACTACTGAAAATGGGCATTCGTGGTGTCTACTGTTTCGATGT 313

QY 315 TCGAGGTTTGGTGTCTCAAGGTGGTTCAACTGAGAGGCGAGCGGCTCAGAAATTTTCTGA 374
Db 314 TCGTGGTTTGGCGCCCAAGGTGGCTTCACTGAGAGGCAAGCTGGCTCTGAAATTTCTGA 373

QY 375 AGACAAGTTTGTCTAAAGTTAAAGTGAAGTCTGTTAGCAAGACCAAGCTTGAGGA 434
Db 374 AGACACGTTTGTTCGAAAAGTTAAATGGAATTTGTTGTCAGCAAGACCAAGCTTGAGG 433

QY 435 TGTTATAGAAAATTCATTGAGAGGCAAGAACTGGAGAGATTGGAGAGCGGCAAGATTTT 494
Db 434 AGTCATTGCCAAGATAATTGAAGAGCAAGACTGTTGAAATAGGTGATGGAAAGATATT 493

QY 495 CTTCGTCCTGTTTTCAGATGTAATTAAGATCGCACTCGTGTGAGCGGGGTGATAAGGCTGA 554
Db 494 CTTCGTCCTGTTTTCAGATGTAATTAAGATCGCACTCGTGTGAGCGGGGTGATAAGGCTGA 553

QY 555 GAGGATGACAGAGGCGGATCTGACATG 582
Db 554 GAGGATGATGGAGGGCATGCTGATATG 581

RESULT 13
BI932123
LOCUS EST552012 tomato flower, 8 mm to preanthesis buds Lycopersicon
DEFINITION esculentum cDNA clone cTOC22G23 5' end, mRNA sequence.
ACCESSION BI932123
VERSION BI932123.1 GI:16246595
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)

```

```

ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 702)
REFERENCE van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
AUTHORS Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,
Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
TITLE Generation of ESTs from tomato flower tissue, buds 8 mm -
preanthesis
JOURNAL Unpublished (2001)
COMMENT Contact: CUGI
Clemson University
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
This clone is available through the Clemson University Genomics
Institute
Seq primer: T3.
Location/Qualifiers
1..702
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cT0C22G23"
/tissue_type="flower"
/dev_stage="buds 8mm to preanthesis"
/clone_lib="tomato flower, 8 mm to preanthesis buds"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Cornell University; sequencing: The
Institute for Genomic Research; flower buds and flowers
were taken from greenhouse plants (4-8 wks old, TA496).
They were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."
ORIGIN
Query Match 46.1%; Score 273.6; DB 12; Length 702;
Best Local Similarity 75.7%; Pred. No. 4.3e-65;
Matches 339; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 135 CGCGGTCAAGCGGTGAAGATATGCCCGTCGTTCTGTGATTAAATGCCCAAGCTCGCC 194
Db 175 CACCTTCAACGCTGTCAAAATGCTCTTCTTCCCGATTATCAGAGCCCAAACTCTCC 234
QY 195 TGACTACATCTCTGATCTAAATCTTACAAGTGGAGCAATCTCAGGCCCTGGCGAGT 254
Db 235 AGATTTTCGTCCTGATGCGAAGTTTACAAGTTGAAGCAATCTAAGACCTTGGAGAT 294
QY 255 CTGCGAAGTTTCTCGCTTTGCTTAAATTTGGTATTCGAGGTGTTACTGTTCTGATGT 314
Db 295 TCACAGGTTTCTCGGCACACTCTGAATATGGGCAATTCGGTGTCTACTGTTGCGAGT 354
QY 315 TCAGAGTTTGTGTGCTCAAGTGTCTCAACTGAGAGCGAGCGGCTCAGAAATTTCTGA 374
Db 355 TCGTGTGTTTGGCGCCCAAGTGTGCTGACTGAGAGCGAGCTGGCTCTGAATTTCTGA 414
QY 375 AGCAAAAGTTTGTGTCTAAAGTTAAGATGAGATGCTGTTAGCAAAAGACAGGTTGAGGA 434
Db 415 AGACAGCTTTGTTGCAAAAGTTAAATGGAATTTGTTGTCAGCAAGACAGGTTGAAGG 474
QY 435 TGTATAGAAAATCATTTGAGGAGCGAAGCACTGGAGAGTTGGAGCGCAAGATTTT 494
Db 475 AGTCATGTCATGATTAATGAGAGGCAAGCACTGGTGAATAGGTGATGGAAGATTT 534
QY 495 CTTGCTCCCTGTTTCAGATGTAATAGAGTCCGCACTGGTGAGCGGGGTGATAGGCTGA 554
Db 535 CTTGACTCCCATCTCCGATGTTAAGAGTTCGCACTGGTGAACGGGAGAAAGGCAGA 594
QY 555 GAGGATGACAGGAGGGCGATCTGCACATG 582
Db 595 GAGGATGACGAGGCGCATGCTGATG 622

```

```

RESULT 14
BI932913
LOCUS EST52802 tomato flower, 8 mm to preanthesis buds Lycopersicon
DEFINITION esculentum cDNA clone cT0C24H17 5' end, mRNA sequence.
ACCESSION BI932913
VERSION BI932913.1 GI:16247385
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 736)
REFERENCE van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J.,
AUTHORS Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M.,
Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
TITLE Generation of ESTs from tomato flower tissue, buds 8 mm -
preanthesis
JOURNAL Unpublished (2001)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
This clone is available through the Clemson University Genomics
Institute
Seq primer: T3.
Location/Qualifiers
1..736
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cT0C24H17"
/tissue_type="flower"
/dev_stage="buds 8mm to preanthesis"
/clone_lib="tomato flower, 8 mm to preanthesis buds"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Cornell University; sequencing: The
Institute for Genomic Research; flower buds and flowers
were taken from greenhouse plants (4-8 wks old, TA496).
They were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."
ORIGIN
Query Match 46.1%; Score 273.6; DB 12; Length 736;
Best Local Similarity 75.7%; Pred. No. 4.3e-65;
Matches 339; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 135 CGCGGTCAAGCGGTGAAGATATGCCCGTCGTTCTGTGATTAAATGCCCAAGCTCGCC 194
Db 137 CACCTTCAACGCTGTCAAAATGCTCTTCTTCCCGATTATCAGAGCCCAAACTCTCC 196
QY 195 TGACTACATCTCTGATCTAAATTTCTACAAGTGGAGCAATTTCTAGGCCCTGGCGAGT 254
Db 197 AGATTTCTGTCCTGATGCGAGTTTACAAGTTTGAAGCAATTTCTAGACCTTGGAGAT 256
QY 255 CTGCGAAGTTTCTCGCTTTGCTTAAATTTGGTATTCGAGGTGTTACTGTTCTGATGT 314
Db 257 TCACAGGTTTCTTCCGCACTACTGAAATGGGCAATTCGTTGTTCTGTTTCCGATGT 316
QY 315 TCAGAGTTTGTGTGCTCAAGTGTCTCAACTGAGAGCGAGCGGCTCAGAAATTTCTGA 374
Db 317 TCGTGTGTTTGGCGCCCAAGTGTGCTGACTGAGAGCGAGCTGGCTCTGATTTCTGA 376
QY 375 AGCAAAAGTTTGTGTCTAAAGTTAAGATGAGATGCTGTTAGCAAAAGACAGGTTGAGGA 434
Db 377 AGACAGCTTTGTTGCAAAAGTTAAATGGAATTTGTTGTCAGCAAGACAGGTTGAAGG 436
QY 435 TGTATAGAAAATCATTTGAGGAGCGCAAGCACTGGAGAGTTGGAGCGCAAGATTTT 494

```

Db 437 AGTCATGCGCATGATTAATGAGAGGCAAGAACTGGTGAATAGGTGATGCGAAGATATT 496
 QY 495 CTTGCTGCGCTGTTTCAGATGTAATAGAGTCCGCATCTGGTCAGCGGGTGATTAAGCTGA 554
 Db 497 CTTGACTCCCATCTCCGATGTTTAAAGAGTTCGCACCTGGTGAACGGGGAGAAAAGGCGA 556
 QY 555 GAGGATGACAGGAGGGCGGATCTGCATG 582
 Db 557 GAGGATGATGGGAGGGCATCTGATATG 584

RESULT 15
 AW738071
 LOCUS EST339498 tomato flower buds, anthesis, Cornell University
 DEFINITION Lycopersicon esculentum cDNA clone cTOD5K11 5', mRNA sequence.
 ACCESSION AW738071
 VERSION AW738071.1 GI:7647016
 KEYWORDS EST.
 SOURCE Lycopersicon esculentum (tomato)
 ORGANISM Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 741)
 van der Hoeven,R.S., Bezzeredes,J.L., Matern,A.L., Holt,I.E.,
 Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
 Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and
 Tanksley,S.D.
 TITLE Generation of ESTs from tomato flower tissue, anthesis
 JOURNAL Unpublished (1999)
 COMMENT Contact: CUGI
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.

FEATURES

source Location/Qualifiers
 1..741
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="cTOD5K11"
 /tissue_type="flower"
 /dev_stages="anthesis"
 /clone_lib="tomato flower buds, anthesis, Cornell University"
 /note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Tanksley; Flower buds and flowers were taken from greenhouse plants (4-8 wks old, TA496). They were immediately frozen in liquid nitrogen and then size-separated while remaining frozen."

ORIGIN

Query Match 46.1%; Score 273.6; DB 1.0; Length 741;
 Best Local Similarity 75.7%; Fred. NO. 4.3e-65;
 Matches 339; Conservative 0; Mismatches 109; Indels 0; Gaps 0;
 QY 135 CGCGGTCAAGCGGTAAAGATATGCCCGCTCGTCTGTGATTAATGCCCAAGCTCGCC 194
 Db 144 CACCTTGAAAGCTGTCAATGCTCTTCTTCGGATTATCAGAGCCCAAACTCTCC 203
 QY 195 TGACTACATTCGTGATGTAATTCACAAAGTGAAGCAATCTCAGGCCCTGCGAGT 254
 Db 204 AGATTTGCTCCCTGATCGAAGTTTACAAAGTTGAAGCAATCTAAGACCTTGAGAA 263
 QY 255 CTCGCAAGTTTCTCGGCTTTCCTAAATTCGATTCGAGGTGTACTGTTTCTGATGT 314
 Db 264 TCAACAGGTTTCTTCGGCACTACTGAAATGGGCATTCGTGGTGACACTGTTTCGATGT 323
 QY 315 TCGAGGTTTGTGCTCAAGGTGGTTCAACTCAGAGGCGAGCGGCGCTCAGAAATTTTCTGA 374

Db 324 TCTGGTTTTGGCGCCCAAGGTGGCTTGACTGAGAGCAAGCTGGCTCTGAAATTTCTCTGA 383
 QY 375 AGACAAGTTTGTGCTAAAGTTTAAGATGGAGATCGTGGTTAGCAAGACCAAGGTTGAGGA 434
 Db 384 AGACACGTTTGTTCAGAAAGTTAAATGGAAATTTGTTGTCAGCAAGACCAAGGTTGAAGG 443
 QY 435 TGTATAGAAAAAATCATTTGAGGAGGCAAGAACTGGAGAGATTGGAGACGCGCAAGATTTT 494
 Db 444 AGTCATTTGCCATGATTAATTGAGAGGCAAGAACTGGTGAATAGGTGATGGAAGAATATT 503
 QY 495 CTTGCTGCGCTGTTTCAGATGTAATAGAGTCCGCATCTGGTCAGCGGGTGATAAGGCTGA 554
 Db 504 CTTGACTCCCATCTCCGATGTTTAAAGATTCCGACTGCTGTAACCGGGAGAAAGGCAGA 563
 QY 555 GAGGATGACAGGAGGGCGGATCTGCATG 582
 Db 564 GAGGATGATGGGAGGGCATCTGATATG 591

Search completed: May 28, 2004, 00:14:55
 Job time : 1377.09 secs